CHARACTERISATION OF HERVS IN AN ASIAN POPULATION: A SINGAPORE PERSPECTIVE

WOO WEE HONG

(M.Sc., NUS)

A THESIS SUBMITTED
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF PATHOLOGY
YONG LOO LIN SCHOOL OF MEDICINE
NATIONAL UNIVERSITY OF SINGAPORE

2014

DECLARATION

I hereby declare that the thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis.

This thesis has also not been submitted for any degree in any university previously.

Woo Wee Hong
6 January 2014

ACKNOWLEDGEMENTS

I would like to express my deepest respect and heartfelt gratitude to my supervisor, Associate Professor Evelyn Koay, for her constant encouragement, guidance and trust throughout my five years of study. I am extremely grateful for all her support and wise counsel in my endeavour to complete this PhD.

I would also like to express my sincere thanks and acknowledgment to my co-supervisor, Dr Leong Sai Mun for his intellectual technical advices and invaluable discussion. His knowledgeable and moral support as well as inspiritment has helped me immensely throughout this long journey.

I am also thankful to my PhD Thesis Advisory and Qualifying Examination Committee members, Associate Professors Chang Chan Fong, Heng Chew Kiat, Vincent Chow and Ng Tze Pin for their time, support and great suggestions.

I wish to thank Dr Eng Chon Boon at the NUH-NUS Tissue Repository for the kind provision of colorectal cell lines. I would also like to thank Mr Lee Hong Kai and Mr Lee Kee Wah at the Molecular Diagnosis Centre, National University Hospital for their expertise advice and assistance on real time thermal cycling and PCR array technology respectively. Many thanks also go to Ms Chua Hui Wen for her great assistance in protein detection technology and purchasing logistics.

I am also grateful to Singapore Polytechnic for providing the financial support for this PhD programme. Special thanks must also go to Dr Adrian

Yeo, Mdm Mah Mon Moey, Dr Ong Chye Sun, Dr Tan Kian Hwa, Dr Tan Tuan Lin, Dr Tan Eng Lee and Mr Edmund Lui for their assistance, support and concern. To all the technical staff of Medical Technology Suite, I thank you for your kind and wonderful technical support.

Last but not least, my deepest appreciation to my wife Siow Yen and two children, Charlene and Markus, for their love and understanding. It is the unwavering support from my wife that makes this journey more meaningful and less distressing. And also, not to forget, the smiling faces of Charlene and Markus that show me the correct spelling of a father in the pursuit of my PhD.

TABLE OF CONTENTS

SUMMARY

Human endogenous retroviruses (HERVs) are a family of viruses within our genome. They represent the trail of successful ancient retroviral infection and constitute about 8% in the human genome. Over the past few years, HERVs have been implicated in physiological functions and pathogenesis of diseases. While the role of HERVs in placental development has been extensively studied, the pathological role of HERVs in diseased states remains to be clearly defined.

Of interest, HERV-H, a family of the HERV families which is characterised by its utilisation of $tRNA^{His}$ at its primer binding site, is selectively expressed in colon cancers but not in normal tissues. Colorectal cancer is the most common cancer for both men and women in Singapore over the past four decades. The average population risk for developing colorectal cancer in Singapore is among the highest in the world. This has led us to further investigate the role of HERV-H in colorectal carcinogenesis. We first demonstrated that the HERV-H gene is not lost from the Singapore population. In fact, HERV-H exists at a high prevalence rate of 89.3% in the population sampled, providing a plausible molecular linkage to the high incidence rate of colorectal cancer in Singapore.

To investigate the role of HERV-H role in neoplastic transformation, we established *in vitro* models to demonstrate the effect of HERV-H on colorectal cancer cells. Our results demonstrated that HERV-H overexpression resulted in the augmentation of cellular proliferative activity, migratory capacity, ability to form cancer stem cell-like colonospheres, and expression

levels of cancer stem cell markers, CD133 and CD44. Furthermore, the gain of decreased serum dependence was most remarkable. In addition it was shown that knock-down of HERV-H significantly lowered cancer cell proliferation, substantiating the oncogenic role of HERV-H in the transforming process. Clearly, the oncogenic role of HERV-H was illuminated with reference to the hallmarks of cancer, namely, limitless replicative potential, tissue invasion & metastasis, self-sufficiency in growth signals and the deregulation of cellular energetic.

Using real-time PCR array, mitogen-activated protein kinase (MAPK), Wnt and p53 pathways were all shown to be implicated in HERV-H-mediated oncogenic signalling. In addition, inflammation-associated players and cytokines were also found to be involved in the HERV-H mediated transformation process. In essence, they represent the evidence that is consistent with the hallmarks of cancer, viz. sustaining proliferative signalling and tumour-promoting inflammation.

Together with previous clinical reports, our data demonstrates the role of HERV-H in colorectal cancer and highlights the molecular basis underlying its involvement. This provides an invaluable resource for rationalizing future therapeutic intervention for colorectal cancers.

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

3D	Three dimensional
AP-1	Activator protein-1
APOBEC3G	Apolipoprotein B mRNA-editing, enzyme-catalytic, polypeptide-like 3G
bp	base pair
cAMP	Cyclic adenosine monophosphate
cDNA	Complementary deoxyribonucleic acid
c-FLIP	cellular FLICE-inhibitory protein
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleoside triphosphate
DsiRNA	Dicer-substrate small interfering ribonucleic acid
EBV	Epstein–Barr virus
ECM	Extracellular matrix
EGF	Epidermal growth factor
EGR1	Early growth response 1
env	Envelope
ERK	Extracellular signal-regulated kinase
ERV	Endogenous retrovirus
FBS	Fetal bovine serum
FGF	Fibroblast growth factor
gag	group-specific antigen
GFP	Green fluorescent protein
GPCR	G protein-coupled receptor
GSEA	Gene set enrichment analysis

HBV	Hepatitis B virus
HCV	Hepatitis C virus
HERV	Human endogenous retroviruse
HHV-8	Human herpesvirus 8
HIV-1	Human immunodeficiency virus type-1
HML	Human MMTV-like sequence
HPV	Human papillomaviruses
HTLV-1	Human T-cell lymphotropic virus type 2
IARC	International Agency for Research on Cancer
ICTV	International Committee on Taxonomy of Viruses
IL	Interleukin
IPA	Ingenuity pathway analysis
JNK	Jun amino-terminal kinase
kDa	Kilodalton
KSHV	Kaposi's sarcoma-associated herpesvirus
LTR	Long terminal repeat
MAPK	Mitogen-activated protein kinase
MEMα	Minimum Essential Medium alpha
MMTV	Mouse mammary tumour virus
mRNA	Messenger ribonucleic acid
NFκB	Nuclear factor kappa-light-chain-enhancer of activated B cells
ORFs	Open reading frames
PBS	Primer binding site
PCR	Polymerase chain reaction

PLZF	Promyelocytic leukemia zinc finger protein
pol	Polymerase
Rb	Retinoblastoma protein
RFU	Relative fluorescence unit
RNA	Ribonucleic acid
RSV	Rous sarcoma virus
SAPK	Stress-activated protein kinase
SD	Standard deviation
siRNA	Small interfering ribonucleic acid
TCF4	Transcription factor 4
TGF-β	Transforming growth factor beta
TRAIL	Tumour necrosis factor-related apoptosis-inducing ligand
TRIM 5α	Tripartite motif 5 alpha
tRNA	Transfer ribonucleic acid
UV	Ultraviolet
Wnt	wingless-type MMTV integration site family

LIST OF PUBLICATIONS

Publication

1. **Woo WH**, Shen L, Leong SM, Koay ES (2014). Prevalence of human endogenous retroviral element associates with Hodgkin's lymphoma incidence rates. Leukemia Research Reports. 3(1):1-3

Meeting Proceedings:

1. **Woo WH**, Koay ES. Distribution of HERV-H in an Asian population. The Inaugural Yong Loo Lin School of Medicine Graduate Scientific Congress. 25 January 2011, Singapore.

2. **Woo WH**, Koay ES, Leong SM (2012). Role of Human Endogenous Retrovirus H in Colorectal Cancer. Annals of the Academy of Medicine. 41(9 Suppl): S213

3. **Woo WH**, Leong SM, Koay ES. The distribution of the endogenous retrovirus HERV-K113 among adolescents. Joint Conference of HGM 2013 and 21st International Congress of Genetics.13 – 18 April 2013, Singapore.

4. **Woo WH**, Leong SM, Koay ES. A cross-sectional study of HERV-H prevalence in the local population of Singapore. Joint Conference of HGM 2013 and 21st International Congress of Genetics.13 – 18 April 2013, Singapore.

5. **Woo WH**, Ang BQ, Woo JY, NG RS, Leong SM, Koay ES. Human endogenous retrovirus H family in colorectal carcinogenesis. 4th Congress of AAMLS and 24th ASM of SAMLS. 2 – 4 Oct 2013, Singapore.

6. **Woo WH**. Tracking the cellular localization of human endogenous retrovirus H family in colorectal cancer cells. 4th Congress of AAMLS and 24th ASM of SAMLS. 2 – 4 Oct 2013, Singapore. (*Oral presentation*)

Introduction

Chapter 1

CHAPTER 1

INTRODUCTION

1.1 Human and viruses

The human race has been continuously exposed to microbial infections throughout their hundreds of millions of years of evolution. Microorganisms like prions, viruses, bacteria, fungi and protozoa, as well as helminths and arthropods can infect human beings and establish a host-parasite relationship. On the other hand, beneficial symbiosis like commensalism and mutualism exist and they play a significant role in moderating the good health of the human digestive system. Still, parasitism remains the choice of life by many different types of microorganisms.

Viruses, for example, are notoriously known to be parasitic (Goering et al., 2008) and have been known to cause many infectious and debilitating diseases such as smallpox (Breman and Henderson, 2002), viral haemorrhagic fever (Paessler and Walker, 2012), poliomyelitis (Racaniello, 2006, Nathanson, 2008), acquired immunodeficiency syndrome (AIDS) (Moir et al., 2011) and influenza (Belshe, 2005). However, another train of thought has suggested otherwise (Fox, 2004, Haig, 2012, Villarreal, 2009, Varela et al., 2009) and put forth the idea of symbiogenetic viruses (Roossinck, 2011, Barr et al., 2013).

1.1.1 Origin of viruses

It has been estimated that there are some 10^{32} virus particles in the biosphere of Earth (Wommack and Colwell, 2000, Bergh et al., 1989). Intriguingly, the sources where these viruses come from remain to be elucidated. An interesting hypothesis known as the Panspermia theory suggested that viruses and other microorganisms originate from outer-space and they rain down upon Earth and bring massive contamination in the earth environment (Raulin-Cerceau et al., 1998). While the academic quest to corroborate the Panspermia theory is not lacking commentary (Wickramasinghe and Trevors, 2013, Demets, 2012), insightful analysis to falsify the theory (Di Giulio, 2010) also exists. This creates an ongoing area of intense research interest in astrobiology (Smith, 2013, Carr et al., 2013).

Although viruses may be structurally similar due to their protein coat, viral genetic materials can only either be DNA or RNA. This observation has led to the generation of three hypotheses (Bandea, 1983). The first hypothesis put forward the idea that viruses were precursors of the earliest cells, indicating that viruses are the relics of pre-cellular life forms. The second hypothesis suggested that viruses were elements of cellular genomes that broke off from the unicellular organisms which were degenerating as a result of viral parasitism. Such viruses were deemed as gene robbers that pick-pocketed the cellular genes. Analyses of viral genomes have revealed the diversity of these genomes and the lack of cellular homology at the sequence level. This has led to the realisation that it is unlikely that all viruses evolved from a single ancestor, the LUCA (the last universal cellular ancestor, or the

Last Universal Common Ancestor) (Woese, 1998, Theobald, 2010). The third hypothesis advocated that viruses originated from fragments of genetic materials that became autonomous and parasitic (Forterre, 2006). While all these hypotheses were interesting and debatable, the origin of viruses remains to be explicitly well-explained (Forterre and Prangishvili, 2009).

1.1.2 Viruses

Viruses are metabolically inert particles that are composed of single-stranded or double-stranded, linear or circular DNA or RNA within a capsid, which is made up of subunits called capsomeres. Human and animal viruses may also possess an additional lipid bilayer membrane or envelope wrapped around the nucleocapsid. Viruses are sized from 30 nm to uncommonly 400 nm. The nomenclature and classification of viruses have been published by the International Committee on Taxonomy of Viruses (ICTV) (International Committee on Taxonomy of Viruses, 2012). To date, the ninth report of the ICTV defines 6 orders, 87 families, 19 subfamilies, 349 genera and 2284 virus and viroid species (King et al., 2011). On the other hand, the Baltimore classification system organises viruses into one of seven classes, based on the type of nucleic acid the virus contains and its mechanism of replication (Baltimore, 1971).

Viruses infect all forms of life. Virophages infect viruses; bacteriophages infect bacteria; mycoviruses infect fungi; plant viruses infect plants. In addition, viruses have also been found to infect invertebrates like

protozoa, worms and arthropods (Wang and Wang, 1991, Bergold, 1953, Felix et al., 2011). Here, the focus will be on human viruses.

Viral infections in humans commence with the entry of virus particles into the host body. The common portal modes of entry include the respiratory tract (via inhaled droplets or saliva), the gastrointestinal tract (in contaminated food or water), skin (typically either by infected insect, animal or human bites or parenteral exposures), genital tract, blood, transplacental and transplants. Following the entry into the human cell, synthesis of viral structural proteins and enzymes as well as replication of the viral genome must take place in order for a successful viral replication cycle to be initiated. Depending on the type of nucleic acid (DNA or RNA) a virus is carrying, synthesis of viral messenger RNA (mRNA) may be preceded by several routes (Table 1.1) so that viral structural or functional proteins can be eventually translated from the monocistronic mRNA and used for its biological activities.

The assembly of a new virus particle is facilitated by the association of the replicated viral genome and the capsomeres. New virus particles are released by budding through the plasma membrane (enveloped virus) or when infected cells are lysed (lytic viruses). Budding often does not damage the cell. Hence the infected cells continue to shed virus particles for long periods, leading to chronic infections that are of important epidemiological consideration. While viruses typically replicate in such manners, viral infections can manifest in an alternative pathway, namely, the lysogenic cycle, in which the proviral DNA becomes incorporated into the host cell chromosome and no progeny virus particles are produced.

Table 1.1. A simplified overview of virus genome replication (Baltimore, 1971)

Virus genome				Mode
DNA	single stranded			ssDNA → dsDNA → mRNA → proteins ↘ ↘ dsDNA → ssDNA ——————> assembly
	double stranded			dsDNA → mRNA → proteins ↘ → dsDNA ——————> assembly
RNA	single stranded	positive sense		+ssRNA ——————> polyprotein ↘ ↘ –ssRNA → +ssRNA → assembly
		negative sense	non-segmented	–ssRNA → +ssRNA ↗ proteins ↘ ↘ ssRNA → assembly
			segmented	
		ambisense		# ambisense RNA → +ssRNA ↘ –ssRNA → +ssRNA
	double stranded			dsRNA → mRNA → proteins → assembly → dsRNA

Proviral DNA continues to replicate along with the cellular DNA in this period of latency. Stimuli are usually needed to trigger a release from latency, leading to a productive infection or lytic cycle. Although the mechanisms regulating the reactivation of latent virus remains incompletely understood to date (Kobayashi et al., 2012), stimulating factors like extreme temperatures, physical trauma, emotional stress and immune suppression are commonly cited.

Integration of proviral DNA may also disturb the regulation of cellular proto-oncogenes since the integrated site may be situated right in, or close to, regulatory regions of the proto-oncogenes and these cellular regulatory sites

may be stimulated by the proviral regulatory elements. Akin to insertional mutagenesis, this leads to the transformation of cells that gives rise to changes in morphology, cellular behaviour and metabolic biochemistry. Controlled growth patterns and contact inhibition are lost, resulting in a state of deranged proliferation of cells and anchorage-independent growth. Conversely, synthesized viral proteins may also be oncogenic due to their interactions with cellular targets such as tumour suppressor protein, disrupting the normal cell cycling profile and facilitating carcinogenesis.

1.2 Human oncogenic viruses

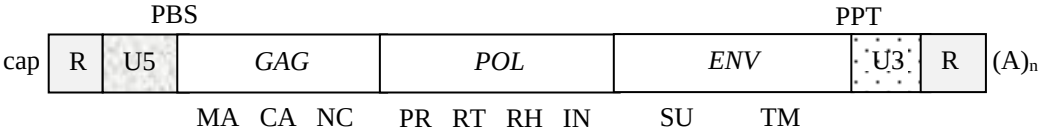
A recent study shows that approximately 16% of the global cancer incidence can be attributed to infections with viruses, bacteria and parasites (de Martel et al., 2012). Specifically, cancer-inducing viruses, or tumour viruses, are the aetiological agents for about 11% of human cancers. There are two classes of human tumour viruses: DNA tumour viruses and RNA tumour viruses which are also known as retroviruses. Currently, there are six tumour viruses which have been classified by the International Agency for Research on Cancer (IARC) as "carcinogenic to humans" (Group I) (Mattock, 2012). These tumour viruses include Epstein–Barr virus (EBV), hepatitis B virus (HBV), human papillomaviruses (HPV) of several types, human T-cell lymphotropic virus type 1 (HTLV-1), hepatitis C virus (HCV), and Kaposi's sarcoma-associated herpesvirus (KSHV), which is also known as human herpesvirus 8 (HHV-8). Although the human immunodeficiency virus type-1 (HIV-1) is also listed as a group 1 cancer-causing agent, it is considered a

cofactor since it primarily augments the carcinogenic action of other viruses like KSHV, EBV and HPV through immunosuppression.

In general, tumour viruses employ two major mechanisms to induce tumours. In direct oncogenesis, the virus infects a progenitor of the clonal tumour cell population, and typically persists in the tumour cells. Indirect oncogenesis occurs when the virus exerts an indirect effect on cell and tissue turnover or on the immune system, predisposing the host to the development of tumour insidiously. The tumour progenitor cell, in this case, may not be invariably infected by the virus. While DNA tumour viruses target the Rb and p53 tumour suppressor gene products and promote cell cycle progression, RNA tumour viruses utilise various means to induce malignancies which include the introduction of viral oncogenes into a cell, the activation of a cellular proto-oncogene and the inactivation of a tumour suppressor gene.

1.2.1 Retroviruses

Retroviruses are RNA viruses which have the ability to convert their viral RNA into a cDNA intermediary by means of a reverse transcriptase and integrating their proviral DNA into the host genome (Varmus, 1988, Bishop, 1978). In general, the retrovirus virion (virus or viral particle) contains two copies of the RNA genome. The two strands of RNA molecules are present as a dimer as a result of complementary base-pairing sequences. The organisation of the RNA genome is depicted in Figure 1.1.



Gene	*GAG*	*g*roup-specific *a*ntig*e*n	
	POL	*pol*ymerase	
	ENV	*env*elope	
Proteins	MA	matrix	lines envelope
	CA	capsid	protects the core; most abundant protein in virus particle
	NC	nucleocapsid	protects the genome; forms the core
	PR	protease	necessary for gag protein cleavage during maturation
	RT	reverse transcriptase	reverse transcribes the RNA genome
	RH	ribonuclease H	degrades RNA and generates PPT primer during reverse transcription
	IN	integrase	required for integration of provirus
	SU	surface glycoprotein	the outer envelope glycoprotein; major viral antigen
	TM	transmembrane glycoprotein	the inner component of the mature envelope glycoprotein
Non-coding sequences	PBS	primer binding site	used by virus to initiate reverse transcription
	PPT	polypurine tract	responsible for initiating (+) strand synthesis during reverse transcription
	R	repeat sequence	"terminally redundant"
	U3	unique sequence at 3' end of genome	forms the 5' end of provirus after reverse transcription; contains promoter elements responsible for provirus transcription
	U5	unique sequence at 5' end of genome	forms the 3' end of provirus after reverse transcription
	$(A)_n$	polyadenylation site	regulates RNA stability and aids in the transport of RNA out of nucleus during provirus transcription

Figure 1.1. Structure of retroviral genome RNA and gene products. RNA genome is capped at the 5' end and polyadenylated at the 3' end. The cap sequence is of type I, m7G5'ppp5''GmpNp. Both the cap and poly(A) tail are attached to the identical short sequences (R). Between these terminal elements are the *gag*, *pol* and *env* genes that encode the structural or functional proteins of the virion.

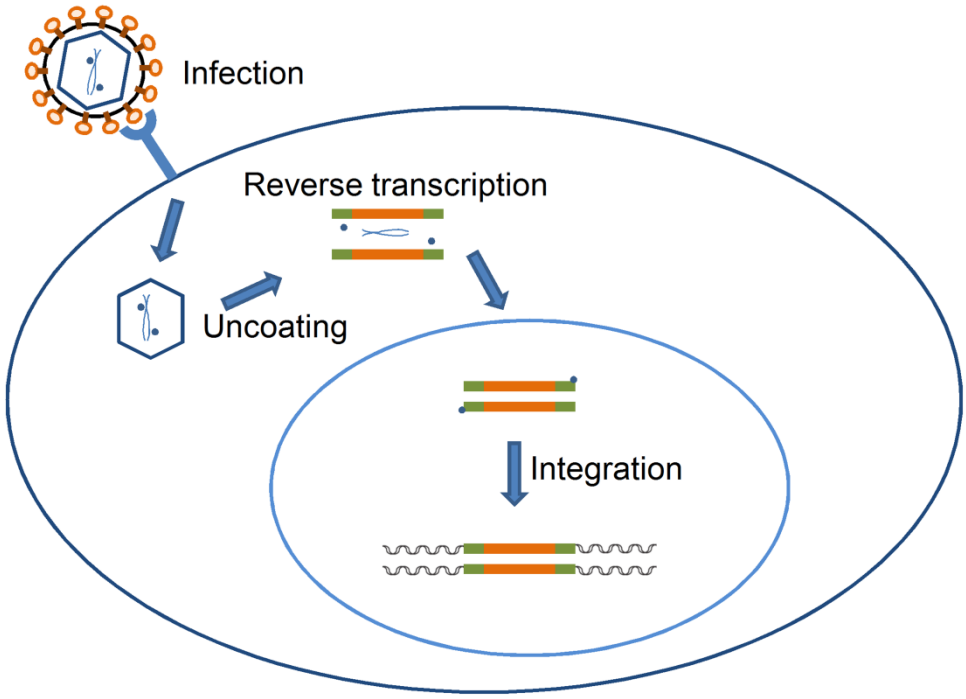


Figure 1.2. How retrovirus integrates its proviral DNA into a host cell. Retrovirus infects its target cell by binding to a special cell surface receptor. Following internalization and uncoating of the retrovirus, RNA is released from the nucleocapsid and is reverse-transcribed into proviral DNA. The provirus is subsequently transported into the nucleus and integrated into the host chromosomal DNA with viral integrase.

The retroviral virion is also packed with enzymes like reverse transcriptase and integrase which are crucial for the replication cycle of retroviruses. When a retrovirus enters the cytoplasm after binding to a specific cell surface receptor, the process of reverse transcription takes place and a double-stranded DNA, termed as provirus is synthesized. The provirus moves into the nucleus and integrates into the host cell genome with the aid of integrase (Figure 1.2). Proviral integration is permanent. With integration, the viral genome is replicated as an integral element of the host genome whenever

the infected cell divides. Thus if proviral integration is present in germ line cells, the provirus can be inherited and may persist as vertical transmission (Longo, 2010). On this note, exogenous retroviruses are considered as infectious while endogenous retroviruses are thought to be non-infectious since they remain as proviral forms.

Historically, exogenous retroviruses were classified under the categories of Types A, B, C and D. These 4 categories were formulated on the basis of viral morphology in negatively stained electron microscopy (Table 1.2). As this classification does not differentiate based on genetic characteristics, such classification is infrequently in use today.

Table 1.2. Historical classification of retroviruses. Retroviruses were once classified based on morphology using negatively stained electron microscopy (Sverdlov, 2005).

Group	Description
A - type	Also known as intracisternal particles. Nonenveloped, immature particles only seen inside cells, thought to result from endogenous retrovirus-like genetic elements
B - type	Enveloped, extracellular particles with a condensed, acentric core and prominent envelope spikes, e.g. Mouse mammary tumour virus
C - type	Same as B-type, but with a central core and barely visible spikes - e.g. most mammalian and avian retroviruses e.g. murine leukaemia virus, avian leukosis virus, human T-cell leukaemia viruses, human immunodeficiency virus
D - type	Usually slightly larger (to 120nm) and spikes less prominent, e.g. Mason-Pfizer monkey virus

Conversely, exogenous retroviruses have also been grouped in three broad groups according to the general details of pathology associated with infection (Table 1.3). These include oncornaviruses (tumour-inducing viruses), lentiviruses (associated with slowly progressing, wasting disease) and spumaviruses ("foamy" viruses associated with persistent infection but without any diseases) (Weiss, 1996).

Table 1.3. Classification of retroviruses based on pathology

Subfamily, Group	Representative	Feature
Oncovirinae (oncogenic viruses)		
Avian leukosis virus	Rous sarcoma virus	Contains *src* oncogene
Mammalian C-type	Abelson leukaemia virus	Contains *abl* oncogene
B-type	Murine mammary tumour virus	Can be endogenous or exogenous
D-type	Mason-Pfizer monkey virus	–
HTLV-BLV	Human T-cell leukaemia virus	Causes T-cell lymphoma and neurologic disease
	Bovine leukaemia virus	Causes lymphoid tumours and lymphosarcoma
Lentivirinae (slow viruses)	HIV-1, HIV-2	Cause AIDS
	Visna virus	Causes lung and brain diseases in sheep
Spumavirinae (foamy viruses)	Simian foamy virus, human foamy virus	Cause no known disease albeit display cytopathogenicity *in vitro*

Based on the differences in morphology and genome organization, the International Committee on Taxonomy of Viruses (ICTV) defines exogenous retroviruses into seven genera (Table 1.4). Under this taxonomy, exogenous retroviral infections usually involve the malignant transformation of an infected cell as well as the induction of an immunodeficiency state that leads

to opportunistic diseases like infections and neoplasms. While exogenous retroviruses mostly infect somatic cells, germ line cells can also be occasionally infected. In such cases, the exogenous virus must be capable of infecting the reproductive tissues in order to infect the germ cell progenitors. The survival of the infected germ cell must take place so that the progeny organism can survive without loss of fitness and ultimately leads to more progeny generations (Blikstad et al., 2008). In this way, the integrated viruses are hereditarily transmitted from one generation to another as stable Mendelian genes and are termed as human endogenous retroviruses (HERVs). Consequently, it is sensible to link high expression levels of HERV elements in reproductive tissues like ovaries, endometrium, testis and placenta (Forsman et al., 2005, Hu et al., 2006) to ancestral introduction and subsequent inheritance of exogenous retroviruses (mode of entry).

Table 1.4. Retrovirus genera using ICTV classification

Genus	Representative	Host	HERV Class[#]
Alpharetrovirus	Rous sarcoma virus	Chickens	**II**
Betaretrovirus	Mouse mammary tumour virus (MMTV)	Mice	**II**
Gammaretrovirus	Murine leukaemia virus (MLV)	Mice	**I**
Deltaretrovirus	Human T cell leukaemia virus type 1 (HTLV-1)	Humans	
Epsilonretrovirus	Walleye dermal sarcoma virus	Fish	
Lentivirus	HIV type Simian immunodeficiency virus Feline immunodeficiency virus	Human Monkeys Cats	
Spumavirus	Simian foamy virus	Monkeys	**III**

[#] HERV classification is not under the purview of ICTV (International Committee on Taxonomy of Viruses)

1.2.2. Endogenous Retroviruses

Human endogenous retroviruses (HERVs) are a family of viruses within our genome and they are the remnants of ancient retroviral infections of the human genome. It is suggested that most HERVs integration took place at least 25 million years ago, before the divergence of apes and Old World monkeys (Shih et al., 1991). Similar to exogenous retroviruses, HERVs are composed of *gag* (group-specific antigen, core protein), *pol* (RNA-dependent DNA polymerase) and *env* (envelope) regions, all of which are flanked by long terminal repeats (LTRs) on both sides (Figure 1.3) (Wilkinson et al., 1994). The classification of HERVs has been convoluted since arbitrary nomenclatures are used by independent investigators, and a constellation of classification criteria like genome complexity, mechanism of replication, type of morphology, copy number and specificity of the tRNA primer binding site (PBS) exists (Urnovitz and Murphy, 1996, Gifford and Tristem, 2003, Bryant et al., 1978, Larsson et al., 1989, Blomberg et al., 2009). In addition, as many HERVs are of fragmentary sequences, there is no consensus as to how this information can be organised into a stable taxonomy (http://ictvonline.org/codeOfVirusClassification_2002.asp).

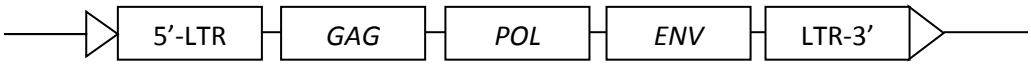


Figure 1.3. Complete structural features of HERV. Arrows indicates repeat sequences (R) generated during integration process; LTR: Long terminal repeat; *GAG* encodes structural retroviral capsid proteins; *POL* encodes enzymes for viral replication, integration, and protein cleavage (polymerase); *ENV* encodes structural retroviral envelope proteins.

Nevertheless, HERVs have been broadly divided into three classes on the basis of sequence similarity to exogenous retroviruses (Tristem, 2000). Class I HERVs show sequence similarity to the gamma-retroviruses (mammalian type C retroviruses) (Table 1.4). They are subdivided into six groups in which three families display homology with murine leukaemia virus (MuLV) and baboon endogenous virus (BaEV) in the highly conserved pol region and the gag and env regions. Members include HERV-H, HERV-I and HERV-R (ERV-9). Class II HERVs show homology to betra-retroviruses and alpha-retroviruses (mammalian type B and D and avian type C retroviruses). Class II HERVs are subdivided into 10 groups. Members include HERV-K and HERV-K(C4) (Dangel et al., 1995). Notably, all class II HERVs are from the HERV-K family, carrying the lysine tRNA specificity and could be further subdivided into type 1 or type 2, on the basis of the presence or absence of a 292 bp segment at the *pol-env* boundary (Barbulescu et al., 1999). Class III HERVs show sequence similarity to the spuma-retroviruses (foamy viruses). Four families are included in this class, namely HERV-S, HERV-L, HERV-U2 and HERV-U3 (Benit et al., 2001).

Commonly, HERVs are further classified into their families using the single letter amino acid code for the tRNA specificity of the primer binding site (PBS). For instance, the HERV-H family contains a PBS similar to $tRNA^{His}$. Currently, over 22 distinct families have been identified in the human genome (Tristem, 2000). The use of the term “family” in describing these HERV lineages is not used in accordance to conventional taxonomic rules as described by the ICTV, since the *Retroviridae* as a whole has been

assigned family status (Blomberg et al., 2009). However, this PBS designation has been widely used in this field (Bannert and Kurth, 2004).

The completion of the Human Genome Project in 2003 revealed the unprecedented information about the genetic make-up in human (Lander et al., 2001). It not only determined the sequences of many genes in human DNA but also intriguingly uncovered the fact that the human genome is almost half occupied – approximately 42% – by transposable elements like the short interspersed elements (SINE), the long-terminal interspersed elements (LINE), retrotransposons and endogenous retroviruses (Prak and Kazazian, 2000).

Interestingly, HERVs are the only elements that possess all the viral coding regions like the *gag*, *pol* and *env* genes that are flanked by long terminal repeats (LTRs), and account for up to 8% of the human genome (Bock and Stoye, 2000) (Figure 1.4). Approximately 400 000 HERVs have been predicted to make up this 8% in the human genome (Ryan, 2004).

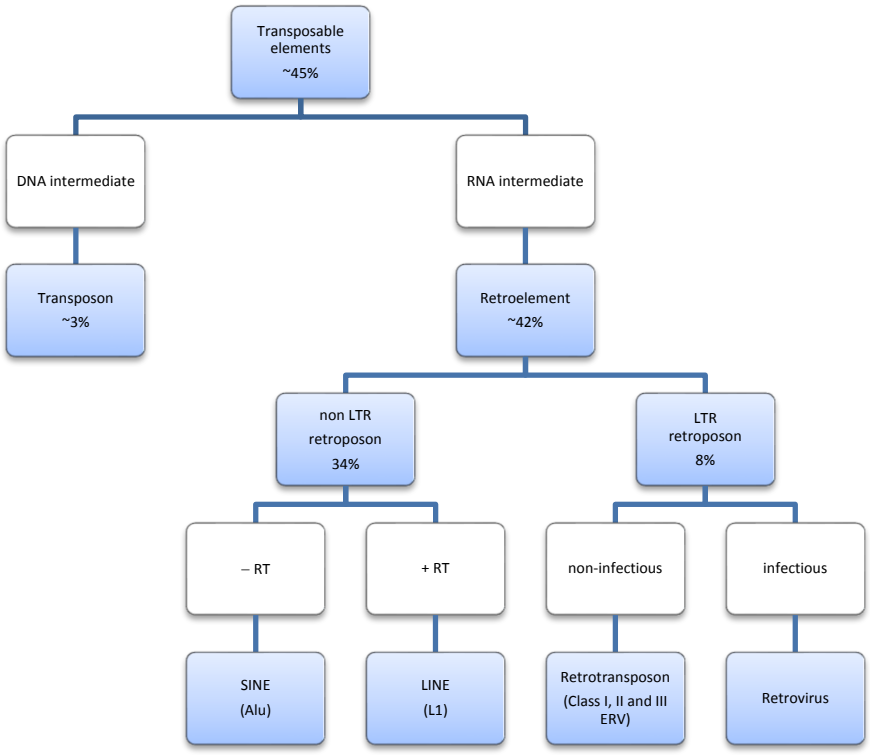


Figure 1.4. Classification of transposable elements. Transposable elements, otherwise known as "jumping genes" are divided into elements that transpose via a DNA intermediate (transposons) or RNA intermediate (retroelements). Retroelements are further subdivided into LTR-containing elements and non-LTR- containing elements. The percentage of each element in the human genome is shown. LTR: long terminal repeat; RT: reverse transcriptase; ERV: endogenous retrovirus; SINE: short interspersed elements; LINE: long interspersed elements; L1: LINE-1; –/+: minus/plus signs.

1.2.3 Various HERV families

HERV was first discovered in 1981 when DNA from human brain specimens was used in Southern blot analysis (Martin et al., 1981). Since then, it is estimated that more than 400 000 copies of HERVs were found and these made up about 50 – 200 HERV families. In view of these data, it was realised

that the constitution of HERVs in the human genome is complex and heterogeneous (Katoh and Kurata, 2013).

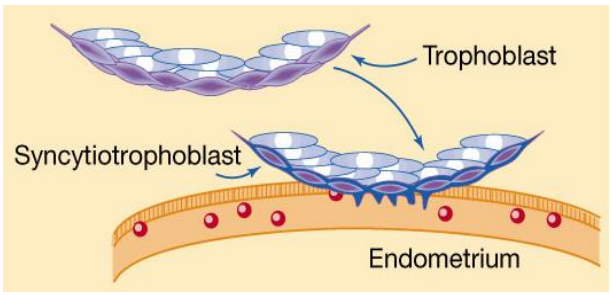
All the 3 classes of HERVs are found in the human genome. HERV-R (O'Connell et al., 1984, Cohen et al., 1985), HERV-E (Steele et al., 1984), HERV-I (Seifarth et al., 2000), HERV-H (Mager and Freeman, 1987, Mager and Henthorn, 1984), ERV-9 (La Mantia et al., 1991), HERV-W (Komurian-Pradel et al., 1999), HERV-T (Werner et al., 1990), HERV-P (Harada et al., 1987) and HERV-FRD (Seifarth et al., 1995) are Class I elements that were reported previously in human specimens. Similarly, Class II elements like the HERV-K superfamily (Andersson et al., 1999, Ono et al., 1986) that comprises groups of HML-1 to 10 (HML stands for *h*uman *M*MTV-*l*ike sequence), and Class III member like HERV-L (Cordonnier et al., 1995, Benit et al., 1999) were also demonstrated in the earlier genetic studies.

1.3 HERVs in health and diseases

Over the past few years, HERVs have been implicated in physiological functions and pathogenesis of diseases (Ryan, 2004). Interestingly, similar to the kind of symbiogenesis as observed for mitochondria in eukaryotes and chloroplasts in plants, HERV-W and HERV-FRD both contribute to the development of the human placenta via their *env* gene products, syncytin 1 (Mi et al., 2000) and syncytin 2 (Blaise et al., 2003) respectively. Syncytins, which possess the cell-cell fusogenic capacity, fuse the trophoblast cells of the placenta and form the syncytiotrophoblasts at the maternal-fetal interface in a

coordinated manner with ERV-3 (also known as HERV-R) (Figure 1.5). Nonetheless, it appears that the role of ERV-3 in the normal physiology of placental development is redundant since 1% of the Caucasian population may have a nonsense mutation in ERV-3 and that does not seem to hinder pregnancy (de Parseval and Heidmann, 1998). Both syncytin-1 and syncytin-2 have been shown to possess immunosuppressive domains. However, only syncytin-2 is found to be functional and pivotal in conferring maternal tolerance to fetal antigens (Mangeney et al., 2007).

(A)



(B)

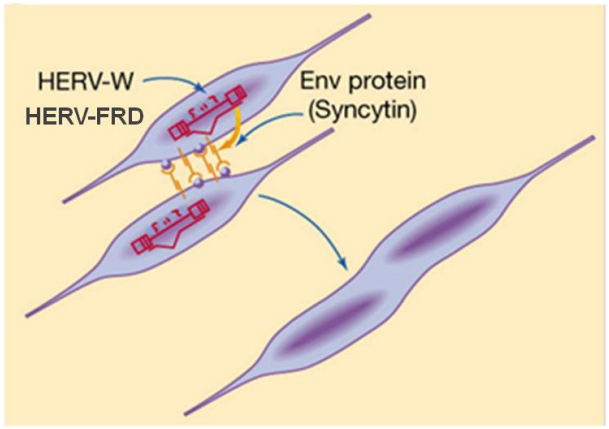


Figure 1.5. Functions of HERV-W and HERV-FRD in human placental development. (A) Fusion of the trophoblast cell layer into the syncytiotrophophoblast. (B) Expression of HERV-W and HERV-FRD *env* proteins, syncytins, mediate trophoblast cell fusion (Adapted and modified from (Stoye and Coffin, 2000)).

Ironically, HERVs might confer some degree of antiviral protection against other exogenous retroviruses, for example, human immunodeficiency virus (HIV). Although HERVs reside in the human genome and therefore their encoded proteins should be regarded as self-antigens, their translated products are antigenic. Overall, HERV proteins are of viral origin essentially. Consequently, immune responses against these HERV proteins via the HERV-specific $CD8^+$ T-cells may help to ameliorate the pathological effects on HIV infected individuals (Garrison et al., 2007, SenGupta et al., 2011, Tandon et al., 2011), although a state of reduced immune activation prior to HIV infection is implicitly favorable (Songok et al., 2012).

Recent evidence has shown that the endogenous retrovirus activity is associated with the pluripotency of embryonic stem cells (Macfarlan et al., 2012), possibly conferring the self-renewal characteristics that embryonic stem cells possess. In addition, demonstration of an abundance of HERV-H transcripts in human embryonic stem cells discloses an intricate relationship between the three important pluripotency transcription factors, namely NANOG, OCT4 and SOX2, and the HERV-H 5' LTR (Santoni et al., 2012) . More studies are needed to provide the role that HERV-H plays in the intricate process of embryonic pluripotency. While the differential expression of various endogenous retroviral elements in early embryos and embryonic stem cells has been shown to link to histone modification machinery and DNA machinery (Rowe and Trono, 2011, Leung and Lorincz, 2012) , how HERV contributes to the development of stem cells remains to be clearly understood.

The role of HERVs in diseased states has always been an interesting topic among many researchers. While HERV-K was frequently thought to play a role in the pathogenesis of diabetes via its superantigen activity (Conrad et al., 1997, Marguerat et al., 2004, Stauffer et al., 2001), controversial findings were observed as well (Badenhoop et al., 1999, Muir et al., 1999, Jaeckel et al., 2002).

HERVs have also been found to be linked with autoimmune diseases (Balada et al., 2010) such as multiple sclerosis (HERV-W, HERV-H) (Petersen et al., 2012, Garson et al., 1998) , rheumatoid arthritis (HERV-W, HERV-K) (Gaudin et al., 2000, Freimanis et al., 2010), systemic lupus erythematosus (Perl et al., 2008, Naito et al., 2003) and Sjogren syndrome (Hishikawa et al., 1997, Price and Venables, 1995). Proposed underlying mechanisms of pathogenesis include molecular mimicry, superantigen activity and LTR-mediated transcriptional activation, all of which putatively lead to the dysregulation of immune responses.

Recently, HERV-K (Huang et al., 2006) and especially HERV-W (Perron et al., 2012a, Leboyer et al., 2013) were reported to have strong associations with psychotic disorders like schizophrenia and bipolar disorders. While the mechanisms remain to be fully elucidated, the insufficiency of neuronal glutamine uptake (Karlsson et al., 2004) and the latent effects of inflammatory neurotoxicity (Leboyer et al., 2011) were proposed. On a similar note, autism spectrum disorder was also found to display a higher expression level of HERV-H and -W (Balestrieri et al., 2012).

1.3.1 HERVs in cancers

More often than not, HERV proteins and transcripts have been detected in various cancerous cells which include teratocarcinoma cell lines, germ cell tumours, melanomas, myeloproliferative disease and cancers of breast, ovary, testis, prostate and colon (Dolei, 2006, Cegolon et al., 2013). While the definitive causal role of HERVs in carcinogenesis remains to be elucidated, it is evident that both Rec and Np9, two accessory viral *env* proteins of HERV-K, are able to upregulate c-*myc* expression by binding and inhibiting c-*myc* gene repressor promyelocytic leukaemia zinc finger protein, leading to a state of increased cell proliferation and reduced apoptosis (Denne et al., 2007). In addition, Rec protein is also found to elevate androgen receptor-mediated transcription, and this may drive tumour induction or promotion in steroid-regulated tissues (Hanke et al., 2013). Conversely, nuclear protein Np9 is able to interact with E3 ubiquitin ligase *L*igand of *N*umb Protein-*X* (LNX), and this interaction may influence the differentiation-proproliferative Notch pathway (Armbruester et al., 2004). Recently, it has been found that Np9 co-activates Wnt/β-catenin, Notch1, ERK and Akt signalling pathways (Chen et al., 2013) that are essential for the survival and proliferation of leukaemia stem/progenitor cells.

Studies have also pointed out that retroviral LTR may be implicated in carcinogenesis (Yu et al., 2013, Katoh and Kurata, 2013). A previous study has reported that hypomethylation of endogenous LTR (LTR of the MaLR family) up-regulates the CSF1R proto-oncogene expression. This in turn

contributes to the pathogenesis of Hodgkin's lymphoma (Lamprecht et al., 2010a, Lamprecht et al., 2010b).

1.3.2 Association between HERV-H and colorectal cancer

HERV-H, a family of the HERVs families, is characterised by its utilisation of histidine (H) tRNA at its primer binding site to initiate reverse transcription. Three independent clinical findings have demonstrated that HERV-H is selectively expressed in colon cancer but not in normal tissues (Liang et al., 2007b, Wentzensen et al., 2007, Alves et al., 2008). Specifically, different transcripts of HERV-H *gag* (Alves et al., 2008) and *env* (Wentzensen et al., 2004) were analysed in these studies. Although the regulatory mechanism of the expression of HERV mRNA remains to be elucidated, it is suggested that transducing agents like hormonal and environmental signals may induce the 5' LTR of HERV and lead to the consequential downstream transcriptional activities. While it may seem that the expression of a HERV sequence in transformed cells could have a contributory role in the development of tumour, it is uncertain whether the increased expression of HERV-H transcripts unerringly precedes the cancer or whether it could be just a result of deranged gene regulation in the tumour cells.

The complete HERV-H sequence on chromosome X was identified and characterised by both Wentzensen et al (2007) and Liang QY et al (2009). A 17-bp sequence in the 5' U3 region was found to possess promoter activity (Liang et al., 2009). Intriguingly, a missing segment on the *env* region of the

HERV-H was reported and the missing *env* region is similar to that of HERV-H/env59, env60 and env62, all of which contain the immunosuppressive domain *in vivo* (Liang et al., 2009, Liang et al., 2007b). Henceforth, should the putative open reading frames (ORFs) contained in this HERV-H transcript be productive, peptides synthesized would be immunogenic, and be able to elicit an immune response that is accompanied by high levels of pro-inflammatory cytokines. Based on the ORFs of HERV-H transcripts, Alves et al (2008) have proposed that a *gag* protein of 93 amino acids (~10.3 kDa) would be generated, suggesting that the induction of immune response could be a likely event.

Conversely, peptides could also be functionally similar to known viral proteins like Rec and Np9. Rec is a 14.5-kDa protein functionally similar to HIV Rev and HTLV Rex, and is known to export unspliced or single spliced viral RNA from the nucleus to the cytoplasm. Together with Np9, a 9-kDa nuclear protein, they bind to the promyelocytic leukaemia zinc finger protein – both a tumour suppressor and transcriptional suppressor of the c-*myc* proto-oncogene – abrogating transcription repression and contributing to increased cell proliferation and reduced apoptosis (Denne et al., 2007).

1.4 Hypothesis and objectives

1.4.1 Rationale

Although a unique constellation of lifestyle factors can reassure a markedly decreased risk of colorectal carcinogenesis (Odegaard et al., 2013),

colorectal cancer remains as the number one cancer in Singapore (Teo and Soo, 2013, Wong and Eu, 2007). There is also a paucity of information regarding the high risk of colorectal cancer development in the local population (Ministry of Health, 2010, National Registry of Diseases Office, 2010, GLOBOCAN 2012, 2012). While diet and nutrition are key modulators in the development of colorectal cancer (Lipkin et al., 1999) and that westernised dietary patterns are commonly associated with colorectal cancers (Randi et al., 2010), it was reported that the high colorectal cancer incidence in the Singapore Chinese population may be associated with insulin resistance, visceral adiposity and physical inactivity instead of dietary patterns (Butler et al., 2008). Additionally, previous studies have reported the involvement of activated c-*myc* and c-*Ki-ras* proto-oncogenes, as well as point mutation of the p53 tumour suppressor gene in the Singapore cohort of colorectal cancer patients (Goh et al., 1996, van Grieken et al., 2013). Nevertheless, the molecular pathogenesis of the high incidence rate of colorectal cancer in Singapore remains to be elucidated.

As aforementioned, HERV-H is found to be differentially expressed in colon cancer but not in adjacent normal tissues. It is of great interest to investigate how HERV-H may be involved in the development of colorectal cancer, which is the aim of the current dissertation.

1.4.2 Hypothesis and objectives

Three independent clinical studies have demonstrated that the expression of HERV-H is selectively expressed in colon cancer but not in normal tissues (Liang et al., 2007b, Wentzensen et al., 2007, Alves et al., 2008). We therefore hypothesized that HERV-H plays a contributory role in the complex process of colorectal carcinogenesis. According to this premise, my present work has the following specific objectives:

(A) To examine the prevalence of HERV-H in the Singapore population

A low distribution frequency or even non-existence of HERV-H in the population genetics of Singapore would sensibly negate the causal role of HERV-H in colorectal carcinogenesis. The existence of HERV-H in the human genome should therefore be verified in the very first instance. By engaging a cross-sectional study that involves the molecular analysis of genomic DNA, the prevalence of HERV-H in the local population can be determined.

(B) To study the effect of HERV-H using *in vitro* models

Using human cell lines as the model of colorectal cancer, the oncogenic potential of HERV-H can be investigated *in vitro*. Overexpression of HERV-H can be done via the transfection of colorectal cancer cell lines with expression vector containing the HERV-H construct. The effect of HERV-H on colorectal cancer cells can then be examined using proliferation assay, scratch assay, invasion assay and cytometric analysis. A knockdown assay using the dicer-substrate small interfering RNA (DsiRNA) is useful in

observing the reversal effect of HERV-H mediating processes. Knockdown assay also serves to substantiate the gain-of-functions via the overexpression study.

(C) To investigate the signalling pathways which are implicated in HERV-H mediating transforming process

Cancer cells are known to acquire the mitogenic signalling they need to sustain proliferation. This is often associated with constitutive activation of signalling pathways. As HERV-H is thought to be involved in the transforming process, it would be useful to find out which cell signalling pathways are linked to the HERV-H mediated oncogenic process. PCR array technology is known to offer an efficient and reproducible approach to differential gene expression profiling (Arikawa et al., 2008, Ning et al., 2008). Use of such a technique enables the profiling of the entire transcriptome in HERV-H mediating transformation process.

An understanding of the role of HERV-H in the pathogenesis of colorectal cancer may provide important diagnostic and therapeutic information. This is useful for the identification of novel prognostic biomarkers in colorectal cancer and the provision of a new strategy for cancer prevention and intervention.

Prevalence of HERV-H in the Singapore population

Chapter 2

CHAPTER 2

PREVALENCE OF HERV-H IN THE SINGAPORE POPULATION

2.1 Introduction

Almost half of the human genome is made up of ancient transposable elements. Distinctively, these transposable elements are segregated into two major classes: DNA and RNA transposons (Finnegan, 1989). While DNA transposons account for about 3% of the human genome and are referred as Class II elements, RNA transposons constitute approximately 42% of the human genome and are termed as Class I elements. In another perspective, approximately 90% of three million transposable elements in the human genome are represented by RNA transposons. RNA transposons are grouped into those without long terminal repeats (LTR), namely the short interspersed elements (SINE), the long-terminal interspersed elements (LINE) and the processed pseudogenes, and those with LTR, viz., the retrotransposons and the endogenous retroviruses (ERVs) (Bannert and Kurth, 2004).

Both DNA and RNA transposons play a role in determining human genotypes and phenotypes on an evolutionary scale and at the individual level (Solyom and Kazazian, 2012). While Class II DNA transposons are able to amplify without an RNA intermediate, Class I retroelements require a reverse-transcribed RNA intermediate to duplicate and insert into new genomic sites (Cordaux and Batzer, 2009).

2.2 Retroelements

Retroelements are widespread in nature. They can be found in vertebrate and invertebrate animals (Boulesteix and Biemont, 2005, Bannert and Kurth, 2004, Eickbush and Furano, 2002), plants (Finnegan, 2012) and fungi (Daboussi and Capy, 2003). Retroelements are organised into LTR-retroposons and non-LTR retroposons. For non-LTR-retroposons, they are further grouped as SINE, LINE and processed pseudogenes. On the other hand, within LTR retroposons, they are further classified into infectious endogenous retroviruses and non-infectious retrotransposons.

Typically, retrotransposons are 5 – 7 kb in length and usually have two open reading frames (ORFs) flanked by LTR on both sides (Figure 2.1) (Lower et al., 1996, Finnegan, 2012). While endogenous retroviruses are capable of retrotransposition from cell to cell via infectious virions, retrotransposons are only able to retrotranspose from one site to another site within the genome in a cell (Finnegan, 2012). It is generally accepted that the differences between retrovirus and retrotransposons are the possession of envelope genes and genomic components needed for making a functional viral capsule (Deininger and Batzer, 2002, Lower et al., 1996).

Infectious LTR retroposons, or specifically the endogenous retroviruses, have long integrated into the human genome. It is estimated that the integration took place at least 30 million years ago (Bannert and Kurth, 2006). In spite of the presence of ORFs, no infectious or autonomously retrotransposing HERVs have been demonstrated.

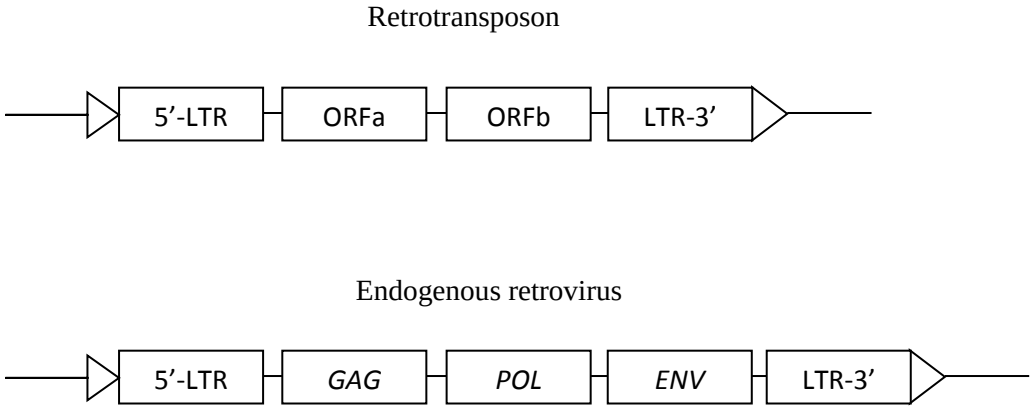


Figure 2.1. Structural features of LTR-retroelements. The LTR-retroelements comprise two members, namely the retrotransposons and the endogenous retroviruses. All LTR-retroelements have a coding region that is flanked by LTRs at both ends. For retrotransposons, the coding region is usually made up of *GAG* and *POL*. Arrows indicates repeat sequences (R) generated during integration process. ORF: open reading frame; LTR: Long terminal repeat; *GAG* encodes structural retroviral capsid proteins; *POL* encodes enzymes for viral replication, integration, and protein cleavage (polymerase); *ENV* encodes structural retroviral envelope proteins.

It was thought that the detrimental consequences of endogenous retroviral activities may have induced host defensive mechanisms to suppress it (Gifford and Tristem, 2003). On the other hand, accumulative effects of mutations or deletions which can cause frame shifts or premature stop codons during host germ line cell division, together with hypermethylation of the promoter sites may have also rendered them defective. Over the years, it was shown that the expression of endogenous retroviruses is actively repressed by the cellular machinery. In addition, the regulation of such endogenous retroviral silencing has been centred upon DNA methylation, histone modifications, and small RNAs (Maksakova et al., 2008).

2.2.1 Human endogenous viruses

Human endogenous retroviruses (HERVs) are organised into 3 classes and further categorised into numerous families. Although HERVs account for about 8% of the human genome, the distribution of HERV families in human genome is heterogeneous (Figure 2.2). Compared to Class II ERVs, both Class I and III ERVs are found to have, even in each of its own class, a higher proportion of the human genome presence. Through analysis of the HERV integration sites, the estimated age of HERVs can be revealed. These findings have offered a likely reason for the percentage difference in human genome presence among the 3 HERVs classes. With reference to Class II HERVs which have been found to integrate within the last 5 million years, both Class I and III HERVs are the oldest groups and they were found to be inserted into the genome at least 25 million years ago (Tristem, 2000).

On the basis that successful integration of proviral DNA into the host cell's chromosomal DNA and endogenization of retroviruses in a reproductive manner are the two necessary processes that must take place during the long period of evolution (Figure 2.3) (Hohn et al., 2013), it should be logical to expect that Class I and III ERVs, being the oldest groups of ERVs, are obviously colonising the genome at large.

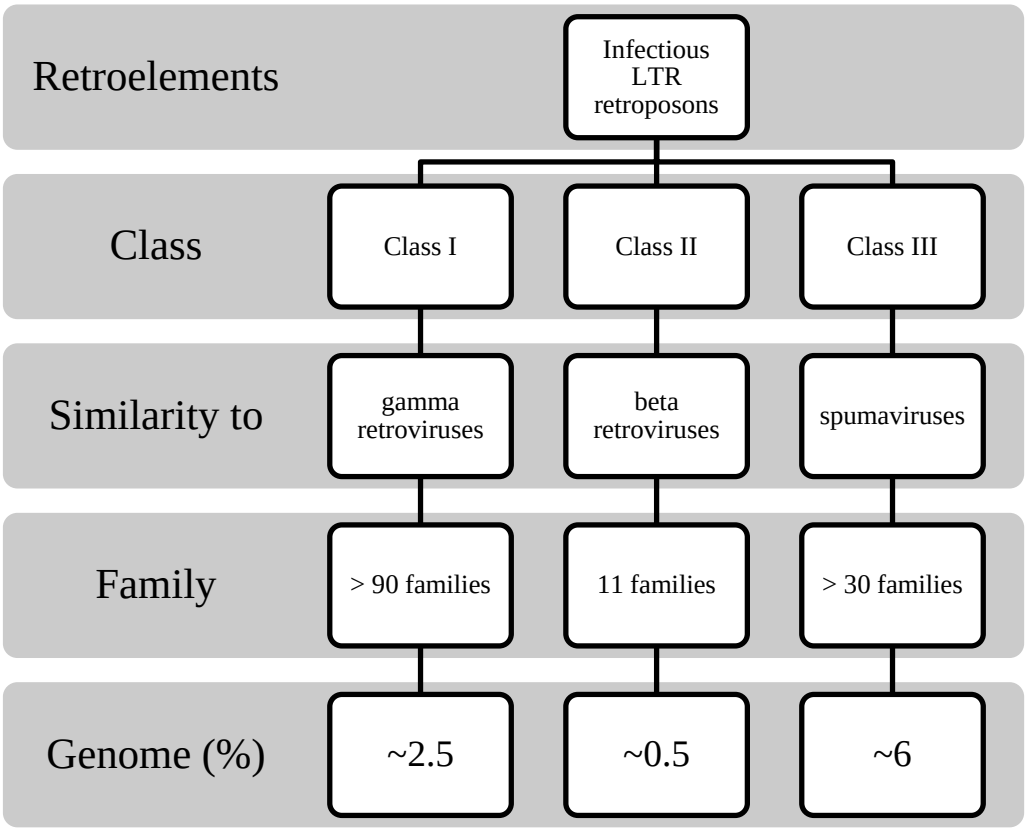


Figure 2.2. LTR-retroelements in human genome. Only infectious LTR-retroposons, which are otherwise known as endogenous retroviruses, is shown here.

The domestication or endogenization of retroviruses is a complex process and takes place during the course of evolution (Figure 2.3). Initially, exogenous retroviruses must reach and infect the germ cell progenitors without being removed by the host's immune system. There are many host factors involved in the resistance to retroviral infection (Takeuchi and Matano, 2008). Retroviral restriction factors like APOBEC3G, a cytidine deaminase which can be encapsidated and hypermutate the new negative cDNA strand during reverse transcription (Cullen, 2006, Yu et al., 2004), TRIM5α, which facilitates the degradation of incoming pre-integration complex via proteasome by ubiquitinylation (Towers, 2005), and BST-2/tetherin, a type 2

integral membrane protein that inhibits the release of fully formed progeny virions (Neil et al., 2008), are some antiviral means that retroviruses must overcome. Following successful provirus integration, the infected germ line progenitors must survive and develop into fertile offspring who can subsequently produce more fertile offspring that carries the proviral DNA. With time, endogenous retroviruses may become fixed in the generations.

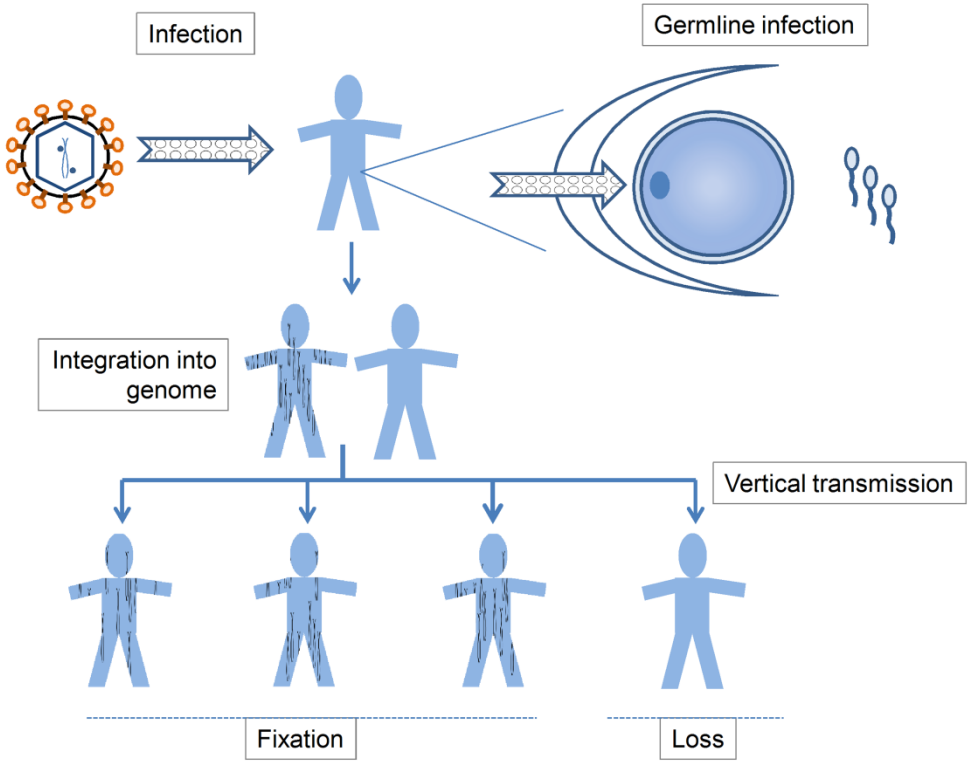


Figure 2.3. The retroviral endogenization process. Exogenous retrovirus is capable of infecting the human cells. For exogenous retrovirus to infect the germ line successfully, it must be able to overcome a myriad of antiretroviral restriction factors. If an infected germ cell containing integrated proviral DNA develops into offspring, the progeny must be fit enough or able to lead to more progeny generations. Thus, the reproductive success of these individuals brings about the transmission of the proviral genome vertically, a process similar to the inheritance of a host gene. During the course of evolution, these proviral genes can either increase in number by the increased frequency of the proviral elements in the population, and thereby leading to the ultimate fixation in the population, or become vanquished by random events or selective pressure against them.

2.2.2 HERV-H

As mentioned in the first chapter, HERVs are implicated in both health and diseases. While HERV-W element has defined its role in health, the causal role of HERVs in diseases remains unclear. With the intent to focus on disease pathology, the HERV-H family is selected and discussed.

HERV-H, a family of the HERVs families, is characterised by its utilisation of histidine (H) tRNA at its the primer binding site (PBS) to initiate reverse transcription. Based on genetic similarity in the *pol* region, HERV-H belongs to Class I of Gammaretrovirus-like. Several studies have shown that the selective expression of HERV-H in colorectal cancer but not in normal tissue is not an isolated event (Liang et al., 2007b, Alves et al., 2008, Wentzensen et al., 2007). While a distinctive pathologic link to HERV-H colorectal carcinogenesis remains to be delineated, there are associated findings that implicate HERV families in cancers like melanoma (Singh et al., 2009), leukaemia (Depil et al., 2002), prostate cancers (Reis et al., 2013) and breast cancers (Wang-Johanning et al., 2008). On the contrary, it is relatively easy to link compelling evidence to exogenous retroviruses in human carcinogenesis. An association of HTLV-1 with adult T-cell leukaemia/lymphoma exemplifies such causal relationship (Cook et al., 2013).

With the premise that HERV-H plays a role in colorectal carcinogenesis, it is of relevance to examine the presence of HERV-H in the genome of an individual, since HERV elements may be present in some humans but not others. While the prevalence of other HERV family members in global populations has been documented, the distribution of HERV-H is not

as well documented. HERV-K113 and HERV-K115 are the two that have been often studied. In one study that comprised 31 genetically diverse humans, 29% (9/31) and 16% (5/31) of individuals tested were found to possess HERV-K113 and HERV-K115 respectively (Turner et al., 2001). Conversely, in a study that involved Asian populations in Taiwan, China and Japan, the average insertion frequency of HERV-K113 was 13% and that of HERV-K115 was 4% (Jha et al., 2009). Interestingly, the insertion frequency of HERV-K113 in a Singapore study (n=120) was 24% (Woo et al., 2013). Taken together, the prevalence of HERV families in various populations is highly variable. Thus, to fill the gap in knowledge about the distribution frequency of HERV-H, we aimed to determine the prevalence of HERV-H in the Singapore population and to draw investigate a plausible molecular linkage to the high incidence rate of colorectal cancer in Singapore.

2.3 Materials and Methods

2.3.1 Study design and subjects

This cross-sectional study was initiated in April 2010 to investigate the prevalence of HERV-H in the population of Singapore. Between April 2010 and Nov 2011, 808 subjects, aged 16 to 60 years, were randomly recruited from the residents in Singapore. The mean age at recruitment was 23.3. Of these, 23 subjects whose genomic materials were classified as indeterminate were excluded from the study, resulting in a total number of 785 subjects.

The study protocol was approved by the Ethics Review Committee of Singapore Polytechnic. All participants gave their written informed consent, and genomic materials were collected after approval by the Ethics Review Committee. Participation in the study was voluntary, anonymous and confidential.

2.3.2 Genotype analysis

Buccal cells were collected using MasterAmp™ buccal brush (Epicentre, Madison, WI). Genomic DNA was isolated from the buccal cells using the ReliaPrep gDNA Tissue Miniprep System (Promega, Madison, WI). Each individual's DNA was tested for the presence of HERV-H by using polymerase chain reaction (PCR) amplifications previously described (Alves et al., 2008).

Briefly, PCR was conducted using the primers HERV-H: forward, 5'-CTT CCC TCC GTG TCT TTA CG-3' and reverse, 5'- AAG ATT AGA CAC ACT CAG CAA CG-3'. The PCR mix, in volume of 50 μl, contained 20 ng of extracted genomic DNA template, 2.0 μM of each primer, 200 μM of each deoxynucleoside triphosphate, 15 mM of MgCl2, and 2.5 units of GoTaq® DNA DNA polymerase (Promega, Foster City, CA). PCR amplification was conducted with 2 min of initial denaturing at 95°C, 35 cycles of 30 seconds at 95°C, 30 seconds at annealing temperature at 60°C, and 60 seconds of elongation at 72°C followed by a 10 min final extension at 72°C. To monitor for reagent contamination, a no template control (NTC) was included in every batch of PCR amplification. A housekeeping gene, β-

actin, was used to validate the true negatives. Primers β-actin: forwards, 5'-CCT TCC TGG GCA TGG AGT CCT G-3' and reverse, 5'-GGA GCA ATG ATC TTG ATC TTC-3' were used. The PCR cycling profile of β-actin followed exactly the same as HERV-H. Amplified PCR products were stored at 4°C until they were analyzed by electrophoresis on 1% agarose gel. The 100-bp ladder (New England Biolabs, USA) was used to verify the product size. The strategy employed in genotype analysis is briefly summarised in Figure 2.4.

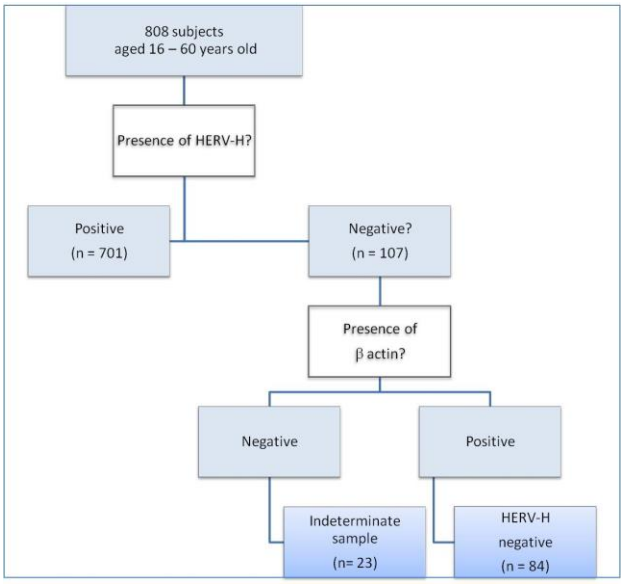


Figure 2.4. Strategy used for the HERV-H genotyping analysis.

2.3.3. Analysis of HERV-H sequence

To examine the magnitude of sequence homology of HERV-H, a specific pair of primers that flanks HERV-H on the either side of chromosome X was designed. The sequence of this pair of primers are: forward, 5'- TTG

GAA ACC TCG CTG GGC TAA AAG C-3' and reverse, 5'-AGA ACT GAT TGA ACG CAA GAA GGC A-3'. A 6942bp amplicon containing the 5462-bp HERV-H (GenBank: EF194101.1) was expected. The sequence matches the region from 2340319 to 2345780 on the human X chromosome (GenBank: NT_167197.1).

Long range PCR amplifications of HERV-H were carried out in 25µl volumes each containing 1X Phusion™ HF Buffer, 200µM of dNTPs, 0.5µM of each primers, 0.5 units of Phusion™ DNA Polymerase (Finnzymes, Espoo, Finland) and 120ng genomic DNA template. Thermal cycler PCR conditions were 98°C for 30 seconds followed by 30 cycles of 98°C for 10 seconds, 64°C for 30 seconds and 72°C for 3 minutes and 15 seconds. Final extension was done at 72°C for 10 minutes. Individual reactions were combined and purified. Purified amplicons were subsequently sequenced using 22 set of sequencing primers (Table 2.1), of which were designed to bind sequentially with overlapping interval onto the various sites of 6942bp HERV-H amplicons. 22 segments were analysed with reference to human chromosome X (GenBank: NT_167197.1). Subsequently, these 22 segments were assembled to form a contiguous sequence with overlapping sequences omitted. The assembled contiguous sequence was analysed with reference to the HERV-H gene in Genbank (EF194101.1). Gene analyses were carried out using BLAST programme (NCBI) or CLUSTAL Omega (EMBL-EBI).

Table 2.1. Sequencing primers for HERV-H

Primer	Sequence	Location of Binding (with reference to GenBank: NT_167197.1)
NEA 0	GAG ATC CAC CCC TGC CCA	2345861
NEA 1	ATC AAT CCT CCG TCC TCC TG	2345614
NEA 2	CAG CCT TCC CTT GGT GTT TA	2345314
NEA 3	TGT CTC TAC TCC TTC TCT GC	2345089
NEA 4	ACC CCT TCC CTC CGT GTC TT	2344789
NEA 5	TTA CAC ATC CGT CCC TTC CT	2344474
NEA 6	CAC TCC TCC ACC CTG TAA TCT TTT T	2344276
NEA 7	GCG TTT AGG CTC TTT TTC AT	2344105
NEA 8	AAA CCC CAG CCA CAT CTC CA	2343825
NEA 9	GAC TCC TTC CCA GAT CTT CT	2343565
NEA 10	ATA CTC TTT TAC GCA CTC CT	2343293
NEA 11	ATC ACC CTT ACC CCG CTC AA	2343036
NEA 12	ATC CTC AAT ACC TCC CTC TA	2342796
NEA 13	ATC TCC CAA ACC TCA ATC CCT TAC A	2342509
NEA 14	ACT ATG CTC AAC TCA CTC TCT ACA	2342301
NEA 15	GCC TAA TCG CCA CAC ACC AG	2342008
NEA 16	CCT CCC TTC CCT ACA CAT CA	2341841
NEA 17	GTC AAA TCA GCC AAG CAG TT	2341587
NEA 18	TTT ACC ACT TTC CCT TCT CA	2341427
NEA 19	TTT CTC CAA GCC ATC ACA GC	2341182
NEA 20	GCT TCT CAA ATC ATC CAA AAC CGT A	2340972
NEA 21	AAG AAG GCA GGA ATG TCA GG	2340711

2.3.4 Statistical methods

Analyses were conducted separately for gender and ethnicity. Prevalence rates and 95% confidence intervals (95% CIs) were calculated using ProMESA software version 1.62 (EpiCentre, Massey University, New Zealand). Odds ratios (ORs) with 95% CIs were calculated using STATGRAPHICS® Centurion XVI version 16.1.05 (Statpoint Technologies, Inc, Virginia, USA) to estimate the association of HERV-H with the variables. Differences between genders were estimated by chi-square test (statistical significance for P <0.05) using STATGRAPHICS® Centurion XVI. All statistical tests were based on 2-sided probability.

2.4 Results

This study aims to examine the prevalence of HERV-H in Singapore. Of 808 subjects recruited, 785 were successfully analysed for HERV-H. The prevalence rates of HERV-H positivity are shown in Table 2.2. 89.9% (381/427) of females and 88.6% (320/361) of males were found to possess HERV-H, making a total of 89.3% HERV-H positivity. There was no statistical difference between males and females ($P > 0.05$).

Table 2.2. HERV-H positivity in sample Singapore population

	HERV-H positive				*P* value
	n	Prevalence Rate (%)	95% CI[#]	Difference (95% CI)	
Total	701	89.3	87.1 - 91.5	N.A.	N.A.
Gender					
Female	381	89.9	87.8 - 92.0	1.3 (-1.76 - 4.36)	0.583
Male	320	88.6	86.4 -90.8	*reference*	
Race/ethnicity					
Chinese	438	86.1	83.7 - 88.5	*reference*	*reference*
Malay	124	95.4	93.9 - 96.9	9.3 (6.47 - 12.1)	0.0035
Indian	121	96.8	95.6 - 98.0	10.7 (7.98 - 13.42)	0.0009
Other*	18	85.7	83.3 - 88.2	0.4 (-3.04 - 3.84)	0.9652

	HERV-H negative				*P* value
	n	Prevalence Rate (%)	95% CI[#]	Difference (95% CI)	
Total	84	10.7	8.5 - 12.9	N.A.	N.A.
Gender					
Female	43	10.1	7.99 - 12.2	1.14 (0.72 - 1.79)	0.583
Male	41	11.4	9.2 - 13.6	1	
Race/ethnicity					
Chinese	71	13.9	11.5 - 16.3	1	*reference*
Malay	6	4.6	3.13 - 6.07	3.35 (1.42 - 7.89)	0.0035
Indian	4	3.2	1.97 -4.43	4.90 (1.76 - 13.7)	0.0009
Other*	3	14.3	11.9 -16.8	0.97 (0.28 - 3.39)	0.9652

*Other (Burmese, Ceylonese, Eurasian, Filipino, Japanese, Javanese, Pakistani, Punjabi, Sikh, Vietnamese). [#]95% CI: 95% confidence interval; calculated based on Singapore Population in 2011 i.e. 5,183,700. Source: Singapore Department of Statistics

Table 2.3 shows the degree of identity at the various sites of HERV-H amplicons from three subjects of different ethnic group. With reference to human chromosome X (GenBank: NT_167197.1), a high level of homology, ranging from 72% to 99% was generally observed. Similarly, when HERV-H sequence of individual subject was compared to the GenBank HERV-H gene (NCBI EF194101.1), a high range of shared identities, between 95 and 99% was observed (Figures 2.5 – 2.7). Table 2.4 indicates relatively good sequence conservation between three ethnic groups, averaging about 98%. To visualise the four sequences simultaneously, sequences were aligned using CLUSTAL Omega. Figure 2.8 reveals a high percentage of sequence identities.

Table 2.3. Overview of HERV-H sequence homology in three ethnic groups. Sequences were aligned with NCBI BLAST program, by using GenBank-derived X chromosome sequence, NT_167197.1 as the query. Sequencing results that were doubtful due to noisy signals in Sanger sequencing are represented by "~".

Alignment of HERV-H Sequences to Human X chromosome (GenBank: NT_167197.1)			
Primer	**Chinese**	**Malay**	**Indian**
NEA 21	2339799 – 2340701 ***Identities:*** *899/904* ***Homology:*** *99%*	2339799 – 2340698 ***Identities:*** *884/885* ***Homology:*** *99%*	~
NEA 20	2339814 – 2340959 ***Identities:*** *1144/1146* ***Homology****: 99%*	2340336 – 2340951 ***Identities:*** *613/618* ***Homology:*** *99%*	2340418 – 2340952 ***Identities:*** *530/535* ***Homology:*** *99%*
NEA 19	2339930 – 2341161 ***Identities:*** *1216/1232* ***Homology:*** *99%*	2340225 – 2341163 ***Identities:*** *935/939* ***Homology:*** *99%*	~

NEA 18	2340171 – 2341412 ***Identities:*** *1231/1242* ***Homology:*** *99%*	2340316 – 2341412 ***Identities:*** *1089/1093* ***Homology:*** *99%*	2340869 – 2341390 ***Identities:*** *511/522* ***Homology:*** *98%*
NEA 17	2340368 – 2341570 ***Identities:*** *1198/1203* ***Homology:*** *99%*	2340467 - 2341571 ***Identities:*** *1088/1111* ***Homology:*** *98%*	~
NEA 16	2340617 – 2341821 ***Identities:*** *1195/1206* ***Homology:*** *99%*	2340555 – 2341829 ***Identities:*** *1239/1267* ***Homology:*** *97%*	2341159 – 2341827 ***Identities:*** *665/691* ***Homology:*** *96%*
NEA 15	2340782 – 2342000 ***Identities:*** *1210/1219* ***Homology:*** *99%*	2341639 – 2342003 ***Identities:*** *351/366* ***Homology:*** *96%*	~
NEA 14	2341065 – 2342292 ***Identities:*** *1216/1228* ***Homology:*** *99%*	2341060 – 2342292 ***Identities:*** *1211/1235* ***Homology:*** *98%*	2341335 - 2342289 ***Identities:*** *951/961* ***Homology:*** *99%*
NEA 13	2341290 – 2342501 ***Identities:*** *1201/1212* ***Homology:*** *99%*	2341379 – 2342501 ***Identities:*** *1114/1123* ***Homology:*** *99%*	~
NEA 12	2341641 – 2342781 ***Identities:*** *1124/1142* ***Homology:*** *98%*	2341971 – 2342773 ***Identities:*** *793/801* ***Homology:*** *99%*	2342417 – 2342770 ***Identities:*** *328/355* ***Homology:*** *92%*
NEA 11	2341796 – 2343023 ***Identities:*** *1216/1228* ***Homology:*** *99%*	2342247 – 2342998 ***Identities:*** *745/752* ***Homology:*** *99%*	~
NEA 10	2342259 – 2343277 ***Identities:*** *1027/1039* ***Homology:*** *99%*	2342026 – 2343277 ***Identities:*** *1225/1252* ***Homology:*** *98%*	2342026 – 2343280 ***Identities:*** *1225/1256* ***Homology:*** *98%*

NEA 9	2342433 – 2343541 ***Identities:*** *1094/1127* ***Homology:*** *97%*	2342282 – 2343541 ***Identities:*** *1236/1262* ***Homology:*** *98%*	2342452 – 2343541 ***Identities:*** *815/1131* ***Homology:*** *72%*
NEA 8	2342566 – 2343818 ***Identities:*** *1226/1253* ***Homology:*** *98%*	2343043 – 2343803 ***Identities:*** *577/768* ***Homology:*** *75%*	2342828 – 2343818 ***Identities:*** *840/1011* ***Homology:*** *83%*
NEA 7	2342843 – 2344096 ***Identities:*** *1217/1254* ***Homology:*** *97%*	2342836 – 2344097 ***Identities:*** *985/1274* ***Homology:*** *77%*	2342838 – 2344097 ***Identities:*** *1094/1270* ***Homology:*** *86%*
NEA 6	2343003 – 2344258 ***Identities:*** *1240/1256* ***Homology:*** *99%*	2342996 – 2344268 ***Identities:*** *1037/1277* ***Homology:*** *81%*	2343218 – 2344272 ***Identities:*** *1015/1060* ***Homology:*** *96%*
NEA 5	2343207 – 2344454 ***Identities:*** *1237/1248* ***Homology:*** *99%*	2343224 – 2344454 ***Identities:****1084/1241* ***Homology:*** *87%*	2343234 – 2344454 ***Identities:*** *1183/1226* ***Homology:*** *96%*
NEA 4	2343521 – 2344780 ***Identities:*** *1237/1260* ***Homology:*** *98%*	2343517 – 2344780 ***Identities:*** *1209/1266* ***Homology:*** *95%*	2343521- 2344780 ***Identities:*** *1228/1260* ***Homology:*** *97%*
NEA 3	2343836 – 2345059 ***Identities:*** *1102/1230* ***Homology:*** *90%*	2343846 – 2345075 ***Identities:*** *1123/1232* ***Homology****: 91%*	2343806 – 2345075 ***Identities:*** *1126/1271* ***Homology:*** *89%*
NEA 2	2344046 – 2345294 ***Identities:*** *1231/1249* ***Homology:*** *99%*	2344042 – 2345297 ***Identities:*** *1234/1256* ***Homology:*** *98%*	2344055 – 2345297 ***Identities:*** *1224/1244* ***Homology:*** *98%*
NEA 1	2344334 – 2345604 ***Identities:*** *1249/1271* ***Homology:*** *98%*	2344330 – 2345598 ***Identities:*** *1253/1271* ***Homology:*** *99%*	2344340 – 2345598 ***Identities:*** *1239/1259* ***Homology:*** *98%*

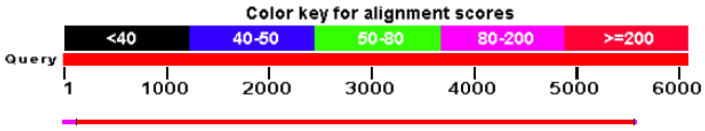
NEA 0	2344913 – 2345847 ***Identities:*** *810/941* ***Homology:*** *86%*	2344617 – 2345847 ***Identities:*** *563/578* ***Homology:*** *97%*	2345259 – 2345832 ***Identities:*** *495/578* ***Homology:*** *86%*

(A)

CCTTTTGATGGTAATTTTCCTTTCCTTTCCCTATATCTATTAAAAGGGCCCACCCCTATCTCCCTTTGCTGACTCTCT
TTTCGGACTCAGCCCGCCTGCACCCAGGTGAAATAAACAGCCTTGTTGCTCACACAAAGCCTGTTTGGTGGTCTCTTC
ACATGGACGCACATGAAATTTGTTGACATTGCTCAGATTGTTAGAACTCCCTTGGGAGATTAATACTCCTTGCTCCTG
CTCTATGTTccccGGGAAAAAACCCCCTACGACCTCAGGTCCTCAGACCGACGAGCCCAAGAAACATCTCACCAATTT
CAAATCTGGTAAGCAGCCTCTTTTTACTCTCTTCTCCAACTTCCCTCACTATCCCTCAACCACTTTCTCCTTTCAATC
TTGGCACTACACTTCAATCTCTCCCTTCTCTTAATTTCAATTCCTTTCATTTTCTGGTAGAGACAAAAGAGACACGTT
TTATCCGTGGACCCAAAACTCCGGCGCTGGTCACAGACTGGGAAGGCAGCCTTCCCTTGGTGTTTAATCATTGCAGGG
ATGCCTCTCTGATTATACACCCACGTTTCAAGGGTGTCAGACCACGCAGGGACAACTGCCTTGGTCCTTCACCCTTAG
CAGCAAGTCCCGCTTTTCTGGGGAAGGGGCAAGTACCCCAACCCCTTCTCTCCGTGTCTCTGTTTTCTCTCGCTTTT
CCGGGGACAGGGCAAGTACCCCAACCCCTTCTCTCCGTGTCTCTACTCCTTCTCTGCTTTTCTGGGGACAGGGCAAGT
ACCCCAACCCCTTCTCTCCTTGTCTCTACCCCTTCTCTGCTTTTCTGGGGCAGGGGCAAGTACCCCTCAACCCCTTCT
CCTTCACCCTTAGCGGTAAGTCCCGCTTTTCTAGGGGGCAGGGGCAAGTACCCCCCAACCCCTTCTCCTTCACCCTTA
GTGGCAAGTCCTGCTTTCCTGGGGCAGGGGCAAGTACCCCTTAACCCCTTCTCCTTCACCCTTAGCAGCAAGTCCCAC
TTTTCTAGGGGGCAAGAAACCCCAAACCCCTTCCCTCCGTGTCTTTACGCTCTCCTTCCTTCACTATGGGCAACCTTC
CATCCTCCATTCCTCCTTCTCCCTTAGCCTGTGTGCTCAAGAACTTAAAACCTCTTCAACTCACACCTGACCTAAAAC
CTAAATGCCTCATTTTCTTCTGCAACACCGCTTGGCCCCAATACAAACTTGACAATGGCTCTAAATGGCCAGAAAATG
GCACTTTCGATTTCTCCATCCTACAAGACCTAAATAATTTTTGTCAAAAAATGGGCAAATGGTCTGAGGTGCCTGATG
TCCAGGCATTCTTTTACACATCCGTCCCTTCCTAGTCTCTGTGCCCAGTGCAACTCGTCCCAAATCTTCCTTCTTTCC
CTCCCACCTGTCCCCTCAGTCCCAACCCCAGGCGTTGCTGAGTGTGTCTAATCTTCCTTTTCTACAAACCCATCTGAC
CTCTCCCCTCCTCTCCACGCCAAGCTAGGTCCCAATTCTTCCTCAGGCTCCACTCCTCCACCCTGTAATCTTTTTATC
ACCTCCCCTCCTCACACCTGGTCCGGCTTACAGTTTCGTTCCGTGACTAGCCCTCCCCCACCTGCCCAGCAATTTACT
CTTAAAAAGGTGGCTGGAGCTAAAGACATAGCCAAGGTTAAAGCTCCCTTTTCCTTATCCCAAATCAAAAACGTTTAG
GCTCTTTTTCATCAAATACAACAACCCAGCCCAATTCATGGCTCGATCGGCAGCAACCTTGATACGCTTTAAAGCCCT
AGAACTTTAAAGGTCAAAACGCCGACTTATTTTCAATACCCATTTTATTACCCAAGCTGCTTCCGAAATTAAAATAAA
ATTCCGACAATTAATTTCCAGGCCTCAAATTCCACCACTTGGATTTAATTAACCTCGCCTTCAAGGTGTACAATAATa
aaaaaaaaaaaGTTGCAATTCCTTGCCTCCACTGTGAGACAAACCCCAGCCACATCTCCAGCAAACAAGAACTTCCAA
ACACCTGAACCGCAGCAGCCAGGCGTTCCTCCAGAACCTCCTCCCACAGGAGCTTGCTACAAGTGCCAGAAATCTGGC
CACCAAGCCAAGAAATGCCTGCAGCCCAGGATTCCTCCTAAGCCGTGTCCCATCTGTGCGGGACCCCATTGGAAATCA
GAATGGTCAACTCACCTGGCAGCGACTCCCAAGAACCCTGGAACTCTGGCCCAAGGTTTCTTGACTGAATCCTTCCCA
GATCTTCTCGGCTTAGCGGCTGAAGACTGACGCTGCCCAATCACCTCGGAAGCCCTGTAGACCATCACAGACACTGAG
CTTTAGGTAACTCTCACAGTGGAGGGTAAGTCCGTCCCCTTCTTAATCAATATGGAGGCTCCCCACTCCACATTACCT
ACTTTTCAAGGGCCTGTTTCCCTTGCCTCCATAACTGCTGTGGGTATTGACGGCCAGGCTTCTAAACCTCTTAAAACT
TCCCAACTCTAGTGCCAACTTAGACAATACTCTTTTACGCACTCCTTTTTAGTTATCCCCACCTGCCCAGTTCCCTTA
TTAGGCCGAGACACTTTAACTAAATTATCTGCTTCCCTGACTACTCCTGGACTACAGCTGCATCTCATTGCCGCCCTT
ATTCCTAATCCAAAGCCTCCTTTGCGTCCTCCTCTTGTATCCCCCCACCTTAACCCACAAGTATAAGATACCTCTACT
CCCTCCTTGGCGACCGATCATGCACCCCTTACCATCTCATTAAAACCTAATCACCCTTACCCCGCTCAATGCCAATAT
CTCATCCCACAGCATGCTTTGAAAGGATTAAAGCCTGTTATCACTTGCCTGCTACAGCATGGGCTCCTAAAACCTATA
AACTCTCCTTACCATTCCCCCATTTTACCTGTCCTAAAACCAGACAAGCCTTACAAGTTAGTTCAGAATCTGCGCCTT
ATCAACCAAATGTTTTGCCTATCCACCCCGTGGTGCCAAACCCGTATACTCTTCTATCCTCAATACCTCCCTCTACTA
CCCATTATTCTGTTCTGGATCTCAAACATGCTTTCTTTACTATTCCTTTGCACCCTTCATCCCAGCCTCTCTTTGCCT
TCACTTAGACTGACCCTGACACCCATTAGGCTCAGCAAATTACCTGGGCTGTACTGCCGCAAGGCTTCACAGACAGCC
CCCATTACTTCAGTCAAGCCCAAATTTCATCCTCATCTGTTACCTATCTCAGCATAATTCTCATAAAAACACAGGTGC
TCTCCCTGCTGATCGTGTCCGATTAATCTCCCAAACCTCAATCCCTTACAAAAGAACAACTCCTTTCCTTCCTAGGCA
TGGTTAGTGTGGTCAGAATTCTTACACAAGAGCCAGGACTGCACCCTGTAGCCTTTCTGTCCAAACAACTTGACCTTC
CTGTTTTAGCCTAGCCATCATGTCTGTGTGCAGCCGCTGCCGCTGCTTTAATACTGTTAGAGGCCCTCAAAATCACAA
ACTATGCTCAACTCACTCTCTACATTTCTCATAACTTCCAAAATCTATTTTCTTCCTCATACCTGACGCATATACTTT
CTGCTCCCCGGCTCCTTCAGCTGTACTCACACTTTCTTAAGTCCCACAATTACCATTGTTCCTGGCCTGGACTTCAAT
CTGGCCTCCCACATTATTCCTGATACCACACCTGACCTCCATGACTGTATCTCTCTGATCCACCTGATATTCACCCCA
TATTTCCTTCTTTCCTGTTCCTCACCCTGATCACGCTTGATTTATTGATGGAAGTTCCACCAGGCCTAATCGCCACAC
ACCAGCAAAGGCAGGTTATGCTATAGTACAAGCCACTAGCCCACCTCTCAGAACCTCTCATTTCCTTTCCATCGTGGA
AATCTATCCTCAAGGAAATAACTTCTCAGTGTTCCATCTGCTATTCTACTACTCCTCAGGGATTATTCAGGCCCCCTC
CCTTCCCTACACATCAGGCTGGAGGATTTGCCCCACCCAGGACTGGCAAATTACCTTTACTCAACATGCCCTGATTCA
GGAAACTAAAATACCTCTTAGTCTAAATAGACACTTTCACTGAATAAAGTAAAGGCCTTTCCTACAGGGTCTGAGAAG
GCCACCACAGTCATTTCTTCCCTTCTGTCAGACACAATTCCTCAGTTTAGCCTTCCCACCTCTATACAGTCTGATAAC
AGACCAGCCTTTATTAGTCAAATCAGCCAAGCAGTTTTTCAGGCTCTTAGTATTCAGTGAAACCTTTATATCCCTTAC
GGTCCTCCTTCTTCAAGAAAAGTAGAACGGACTAAAGGTCTTTTAAAAACACACCTCACCAAGCTCAGCCACCAGCTT
AAAAAGGACTGGACAATACTTTTACCACTTTCCCTTCTCAGAAATCAGGCCTGTCCTCAGAATGCTACAAGGTACATC
CCATTTAAGCTCCTGTATAGACGCTCCTTTTTATTAGGCCCCAGTCTCATTCCAGACACCAGACCAACTTAGACTGTG
CCCCAAAAAACTTGTCATCCCTACTATCTTCTGTCTAGTCATACTCCTATTCACCGTTCTCAACTACTCATACATGCC
CTGCTCTTGTTTACACTGCCGGTTTACACTGTTTCTCCAAGCCATCACAGCTGATATCTCCTGGTGCTATCCCAAACT
GCCACTCTTAACTCTTGAAGTAAATAATCTTTGCTGGCAGGACTATGCTGAATCTCCTTAGGCACTCTCTAATCAGAT
GTCCTAGGTCCTCCCAATTCTTAGACATTTAATACCCATTTTTCTCCTCCTTTTATTCGGACCTTGTATCTTCCATTT
AGCTTCTCAAATCATCCAAAACCGTATCCAGGCCATCACCAATCATTCTATACGACAAATGTTTCTTCTAACATCCCC
ATGATATCACCCCTTACCACAAGACCTCCCTTCAGCTTAATCTCTCCCACTCTAGGCTCCCACGCCGCCCCTAATCCC
GCTTGAAGCAGCCCTGAGAAACATCGCCCATTATCTCTCCATACCACCCCCCAAAAATTTTCGCTGCCCCAACACTTC
AACACTATTTTGTTTTATTTGTCTTATTAATATAAGAAGGCAGGAATGTCAGGCCTCTGAGCCCAGGCCAGGCCATCG
CATCCCCTGTGACTTGCACGTATACATCCAGATGGCCTAAAGTAACTGAAGATCCACAAAAGAAGTAAAAACAGCCTT
AACTGATGACATTCCAACATTGTGATTTGTTCCTGCCCCACCCTAACTGATAAATGTACTTTGTAATCTCCCCCACCC
TTAAGAAGGTCCTTTGTAATTCTCCCCACCCTTGAGAGTGTACTTTGTGAGATCCACACCTGCCCACCAGAGAACAAA
CCCCCTTTGACTGTAATTTTCCATTACCTTCCCTAATCCTATAAAACGGCCCCACCCCATCTCCCTTTGCTGACTCTC
TTTTCGGACTCAGCCCGCCTGCACCCAGGTGAAATAAACAGCCTTGTTGCTCACACAAAGCCTGTTTGGTGGTCTCTT
CACACGGACGCGCGTGAAAGTCATTAACAATGTTCAGTATTTTAAAGAGTTCCCTCAATAAATAATACTCATTGAACC
AACTCGATTTTTGaaaaaaaaaGGAAGAGTCCACTGATCGTGAAAACCTCCCTTGACTTTCTTGTCTAAAAACTTTTC
AGATAGACTAGAACAAAGTTCTACAGCTAAAATGAACATAAAATCACCTGACCTATTTGCTTAACCTGCAACCACCCA

```
AGGGGTTCACCTTGCCTGCTACCTAGACAGAGCCGATTCATTAAGACGGGAATTGCAATAGAGAAAGAGTAATTCGTG
CAGAGCTGGCTGTGCGGGAGATCCGAGTTTTATTATTACTCAAATCAGTCTCCCCGAGCATTCAGGGAGCAGAGCTGT
TAAGGATAACTTGGTGGGTCGGGAGAAGCCAGTGAGTCAGGAGTGCTGATTGGTCAGAGATGAAATCATAGGGAGTCA
GAGCTGCCTTCTTGCG-TCAATCCAGTTCT
```

(B)



Description	Max score	Total score	Query cover	E value	Max ident	Accession
Homo sapiens endogenous virus HERV-H mRNA	9625	10727	92%	0.0	99%	EF194101.1

Range 1: 1 to 5462 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
9625 bits(5212)	0.0	5385/5466(99%)	22/5466(0%)	Plus/Plus

Figure 2.5. Definition of a HERV-H sequence from a Chinese subject. (A) Assembly of a contiguous sequence of 6 kb HERV-H. (B) Alignment scores from NCBI BLAST.

(A)

```
CCCCTTGAAGGTATTTTCATTTCCTTTCCATAACCTATTAAAAGGccccccccccTATCTCCCTTTGCTGACTCTCTTT
TCGGACTCAGCCCGCCTGCACCCAGGTGAAATAAACAGCCTTGTTGCTCACACAAAGCCTGTTTGGTGGTCTCTTCAC
ATGGACGCACATGAAATTTGTTGACATTGCTCAGATTTTTAGAACTCCCTTGGGAAATTAATCCTCATTCCTCCTGCT
CTTTGTTTTGAGAAAAAACCACCTACGACCTCAGGTCCTCAGACCGACGAGCCCAAGAAACATCTCACCAATTTCAAA
TCTGGTAAGCAGCCTCTTTTTACTCTCTTCTCCAACTTCCCTCACTATCCCTCAACCACTTTCTCCTTTCAATCTTGG
CACTACACTTCAATCTCTCCCTTCTCTTAATTTCAATTCCTTTCATTTTCTGGTAGAGACAAAAGAGACACGTTTTAT
CCGTGGACCCAAAACTCCGGCGCTGGTCACAGACTGGGAAGGCAGCCTTCCCTTGGTGTTTAATCATTGCAGGGATGC
CTCTCTGATTATACACCCACGTTTCAAGGGTGTCAGACCACGCAGGGACAACTGCCTTGGTCCTTCACCCTTAGCAGC
AAGTCCCGCTTTTCTGGGGAAGGGGCAAGTACCCCAACCCCTTCTCTCCGTGTCTCTACCCCTTCTCTGCTTTTCCGG
GGACAGGGCAAGTACCCCAACCCCTTCTCTCCGTGTCTCTACTCCTTCTCTGCTTTTCTGGGGACAGGGCAAGTACCC
CAACCCCTTCTCTCCTTGTCTCTACCCCTTCTCTGCTTTTCTGGGGCAGGGGCAAGTACCCCTCAACCCCTTCTCCTT
CACCCTTAGCGGTAAGTCCCGCTTTTCTAGGGGGCAGGGGCAAGTACCCCCCAACCCCTTCTCCTTCACCCTTAGTGG
CAAGTCCTGCTTTCCTGGGGCAGGGGCAAGTACCCCTTAACCCCTTCTCCTTCACCCTTAGCAGCAAGTCCCACTTTT
CTAGGGGGCAAGAAACCCCAAACCCCTTCCCTCCGTGTCTTTACGCTCTCTTTTCTCTGGGTTTGCTTCCTTCACTAT
GGGCAACCTTCCATCCTCCATTCCTCCTTCTCCCTTAGCCTGTGTGCTCAAGAACTTAAAACCTCTTCAACTCACACC
TGACCTAAAACCTAAATGCCTCATTTTCTTCTGCAACACCGCTTGGCCCCAATACAAACTTGACAATGGCTCTAAATG
GCCAGAAAATGGCACTTTCGATTTCTCCATCCTACAAGACCTAAATAATTTTTGTCAAAAAATGGGCAAATGGTCTGA
GGTGCCTGATGTCCAGGCATTCTTTTACACATCCGTCCCTTCCTAGTCTCTGTGCCCAGTGCAACTCGTCCCAAATCT
TCCTTCTTTCCCTCCCACCTGTCCCCTCAGTCCCAACCCCAGGCGTTGCTGAGTGTGTCTAATCTTCCTTTTCTACAA
ACCCATCTGACCTCTCCCCTCCTCTCCACGCCAAGCTAGGTCCCAATTCTTCCTCAGGCTCCACTCCTCCACCCTGTA
ATCTTTTTATCACCTCCCCTCCTCACACCTGGTCCGGCTTACAGTTTCGTTCCGTGACTAGCCCTCCCCCACCTGCCC
AGCAATTTACTCTTAAAAAGGTGGCTGGAGCTAAAGACATAGTCAAGGTTAATGCTCCTTTTTCTTTATCCCAAATCA
GAAGCGTTTAGGCTCTTTTTCATCAAATATAAAAACCCAGCCCAGTTCATGGCTCGTTCGGCAGCAACCCTGAGACGC
TTTACAGCCCTAGACCCTAAAAGGTCAAAAGGCCGTCTTATTCTCAATATACATTTTATTACCCAATCTGCTCCCGAT
ATTAAATAAAACTCCAAAAATTAAATTCCGGCCCTCAAACTCCACAACAGGATTTAATTAACCTCGCCTTCAAGGTGT
ACAATAATaaaaaaaaaaaaTTGGAATTTCTTTGCTCCACTGGGaaaaaaaaCCCAACCACCTTTCCCGGaaaaaaaa
aCTTTCAAACACCTTAACCGCAGCAGCCAGGCGTTCCTCCAGAACCTCCTCCCACAGGAGCTTGCTACAAGTGCCAGA
AATCTGGCCACCAAGCCAAGAAATGCCTGCAGCCCAGGATTCCTCCTAAGCCGTGTCCCATCTGTGCGGGACCCCATT
GGAAATCAGACTGTTCAACTCACCTGGCAGCGACTCCCAGAGACCCTGGAACTCTGGCCCAGGGTTTTTTAATGGATT
CTTTCCCAAATTTTTTCGGTTAACGGGTGAAGAATGGACGCTGCCCAATCACCTCGGAGCCCTGTAGACCATCACAGA
CACTGAGCTTTAGGTAACTCTCACAGTGGAGGGTAAGTCCGTCCCCTTCTTAATCAATATGGAGGCTCCCCACTCCAC
ATTACCTACTTTTCAAGGGCCTGTTTCCCTTGCCTCCATAACTGCTGTGGGTATTGACGGCCAGGCTTCTAAACCTCT
TAAAACTTCCCAACTCTAGTGCCAACTTAGACAATACTCTTTTACGCACTCCTTTTTAGTTATCCCCACCTGCCCAGT
TCCCTTATTAGGCCGAGACACTTTAACTAAATTATCTGCTTCCCTGACTACTCCTGGACTACAGCTGCATCTCATTGC
CGCCCTTATTCCTAATCCAAAGCCTCCTTTGCGTCCTCCTCTTGTATCCCCCCACCTTAACCCACAAGTATAAGATAC
CTCTACTCCCTCCTTGGCGACCGATCATGCACCCCTTACCATCTCATTAAAACCTAATCACCCTTACCCCGCTCAATG
CCAATATCTCATCCCACAGCATGCTTTGAAAGGATTAAAGCCTGTTATCACTTGCCTGCTACAGCATGGGCTCCTAAA
ACCTATAAACTCTCCTTACCATTCCCCCATTTTACCTGTCCTAAAACCAGACAAGCCTTACAAGTTAGTTCAGAATCT
```

GCGCCTTATCAACCAAATGTTTTGCCTATCCACCCCGTGGTGCCAAACCCATATACTCTTCTATCCTCAATACCTCCC
TCTACTACCCATTATTCTGTTCTGGATCTCAAACATGCTTTCTTTACTATTCCTTTGCACCCTTCATCCCAGCCTCTC
TTTGCCTTCACTTAGACTGACCCTGACACCCATTAGGCTCAGCAAATTACCTGGGCTGTACTGCCGCAAGGCTTCACA
GACAGCCCCCATTACTTCAGTCAAGCCCAAATTTCATCCTCATCTGTTACCTATCTCAGCATAATTCTCATAAAAACA
CAGGTGCTCTCCCTGCTGATCGTGTCCGATTAATCTCCCAAACCTCAATCCCTTACAAAAGAACAACTCCTTTCCTTC
CTAGGCATGGTTAGTGTGGTCAGAATTCTTACACAAGAGCCAGGACTGCACCCTGTAGCCTTTCTGTCCAAACAACTT
GACCTTCCTGTTTTAGCCTAGCCATCATGTCTGTGTGCAGCCGCTGCCGCTGCTTTAATACTGTTAGAGGCCCTCAAA
ATCACAAACTATGCTCAACTCACTCTCTACATTTCTCATAACTTCCAAAATCTATTTTCTTCCTCATACCTGACGCAT
ATACTTTCTGCTCCCCGGCTCCTTCAGCTGTACTCACACTTTCTTAAGTCCCACAATTACCATTGTTCCTGGCCTGGA
CTTCAATCTGGCCTCCCACATTATTCCTGATACCACACCTGACCTCCATGACTGTATCTCTCTGATCCACCTGATATT
CACCCCATATTTCCTTCTTTCCTGTTCCTCACCCTGATCACGCTTGATTTATTGATGGAAGTTCCACCAGGCCTAATC
GCCACACACCAGCAAAGGCAGGTTATGCTATAGTACAAGCCACTAGCCCACCTCTCAGAACCTCTCATTTCCTTTCCA
TCGTGGAAATCTATCCTCAAGGAAATAACTTCTCAGTGTTCCATCTGCTATTCTACTACTCCTCAGGGATTATTCAGG
CCCCCTCCCTTCCCTACACATCAGGCTGGAGGATTTGCCCCACCCAGGACTGGCAAATTACCTTTACTCAACATGTCC
TGAGTCAGGAAACTAAAATACCTCTTAGTCTAAATAGACACTTTCACTGAATAAAGTAAAGGCCTTTCCTACAGGGTC
TGAGAAGGCCACCACAGTCATTTCTTCCCTTCTGTCAGACACAATTCCTCAGTTTAGCCTTCCCACCTCTATACAGTC
TGATAACAGACCAGCCTTTATTAGTCAAATCAGCCAAGCAGTTTTTCAGGCTCTTAGTATTCAGTGAAACCTTTATAT
CCCTTACGGTCCTCCTTCTTCAAGAAAAGTAGAACGGACTAAAGGTCTTTTAAAAACACACCTCACCAAGCTCAGCCA
CCAGCTTAAAAAGGACTGGACAATACTTTTACCACTTTCCCTTCTCAGAAATCAGGCCTGTCCTCAGAATGCTACAAG
GTACATCCCATTTAAGCTCCTGTATAGACGCTCCTTTTTATTAGGCCCCAGTCTCATTCCAGACACCAGACCAACTTA
GACTGTGCCCCAAAAAACTTGTCATCCCTACTATCTTCTGTCTAGTCATACTCCTATTCACCGTTCTCAACTACTCAT
ACATGCCCTGCTCTTGTTTACACTGCCGGTTTACACTGTTTCTCCAAGCCATCACAGCTGATATCTCCTGGTGCTATC
CCCAAACTGCCACTCTTAACTCTTGAAGTAAATAATCTTTGCTGGCAGGACTATGCTGAATCTCCTTAGGCACTCTCT
AATCAGATGTCCTAGGTCCTCCCAATTCTTAGACATTTAATACCCATTTTTCTCCTCCTTTTATTCGGACCTTGTATC
TTCCATTTAGCTTCTCAAATCATCCAAAACCGTATCCAGGCCATCACCAATCATTCTATACGACAAATGTTTCTTCTA
ACATCCCCATGATATCACCCCTTACCACAAGACCTCCCTTCAGCTTAATCTCTCCCACTCTAGGCTCCCACGCCGCCC
CTAATCCCGCTTGAAGCAGCCCTGAGAAACATCGCCCATTATCTCTCCATACCACCCCCCAAAAATTTTCGCTGCCCC
AACACTTCAACACTATTTTGTTTTATTTGTCTTATTAATATAAGAAGGCAGGAATGTCAGGCCTCTGAGCCCAGGCCA
GGCCATCGCATCCCCTGTGACTTGCACGTATACATCCAGATGGCCTAAAGTAACTGAAGATCCACAAAAGAAGTAAAA
ACAGCCTTAACTGATGACATTCCAACATTGTGATTTGTTCCTGCCCCACCCTAACTGATAAATGTACTTTGTAATCTC
CCCCACCCTTAAGAAGGTCCTTTGTAATTCTCCCCACCCTTGAGAGTGTACTTTGTGAGATCCACACCTGCCCACCAG
AGAACAAACCCCCTTTGACTGTAATTTTCCATTACCTTCCCTAATCCTATAAAACGGCCCCACCCCATCTCCCTTTGC
TGACTCTCTTTTCGGACTCAGCCCGCCTGCACCCAGGTGAAATAAACAGCCTTGTTGCTCACACAAAGCCTGTTTGGT
GGTCTCTTCACACGGACGCGCGTGAAAGTCATTAACAATGTTCAGTATTTTAAAGAGTTCCCTCAATAAATAATACTC
ATTGAACCAACTCGATTTTTGaaaaaaaaaGGAAGAGTCCACTGATCGTGAAAACCTCCCTTGACTTTCTTGTCTAAA
AACTTTTCAGATAGACTAGAACAAAGTTCTACAGCTAAAATGAACATAAAATCACCTGACCTATTTGCTTAACCTGCA
ACCACCCAAGGGGTTCACCTTGCCTGCTACCTAGACAGAGCCGATTCATTAAGACGGGAATTGCAATAGAGAAAGAGT
AATTCGTGCAGAGCTGGCTGTGCGGGAGATCCGAGTTTTATTATTACTCAAATCAGTCTCCCCGAGCATTCAGGGAGC
AGAGCTGTTAAGGATAACTTGGTGGGTCGGGAGAAGCCAGTGAGTCAGGAGTGCTGATTGGTCAGAGATGAAATCATA
GGGAGTCAGAGCTGCCTTCTTGCGTTCAATTCAGTTCT

(B)

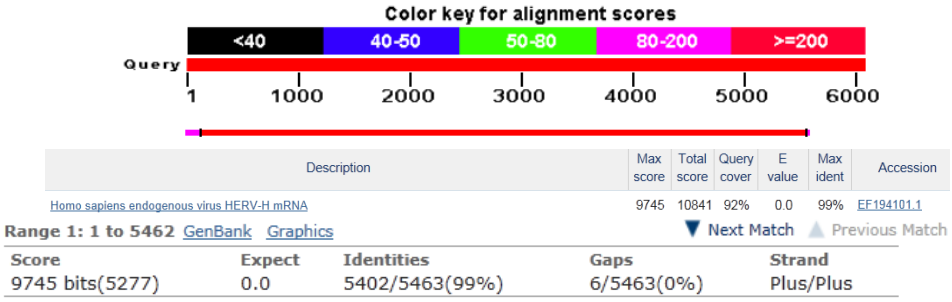


Figure 2.6. Definition of a HERV-H sequence from a Malay subject. (A) Assembly of a contiguous sequence of 6 kb HERV-H (B) Alignment scores from NCBI BLAST.

(A)

TTTNCNTTNCCTTNCCCTATCCTATTAAAAGGCCCCACCCCTATCTCCCTTTGCTGACTCTCTTTTCGGACTCAGCCC
GCCTGCACCCAGGTGAAATAAACAGCCTTGTTGCTCACACAAAGCCTGTTTGGTGGTCTCTTCACATGGACGCACATG
AAATTTGTTGACATGGCTCAGATCGGGGGACCTCCCCTGGGAGATCAATCCTCCTTGCTCCTGCTCTTTGCTTGaaaa

GAAAAATCACCTACGACCTCAGGTCCTCAGACCGACGAGCCCAAGAAACATCTCACCAATTTCAAATCTGGTAAGCAG
CCTCTTTTTACTCTCTTCTCCAACTTCCCTCACTATCCCTCAACCACTTTCTCCTTTCAATCTTGGCACTACACTTCA
ATCTCTCCCTTCTCTTAATTTCAATTCCTTTCATTTTCTGGTAGAGACAAAAGAGACACGTTTTATCCGTGGACCCAA
AACTCCGGCGCTGGTCACAGACTGGGAAGGCAGCCTTCCCTTGGTGTTTAATCATTGCAGGGATGCCTCTCTGATTAT
ACACCCACGTTTCAAGGGTGTCAGACCACGCAGGGACAACTGCCTTGGTCCTTCACCCTTAGCAGCAAGTCCCGCTTT
TCTGGGGAAGGGGCAAGTACCCCAACCCCTTCTCTCCGTGTCTCTACCCCTTCTCTGCTTTTCCGGGGACAGGGCAAG
TACCCCAACCCCTTCTCTCCGTGTCTCTACTCCTTCTCTGCTTTTCTGGGGACAGGGCAAGTACCCCAACCCCTTCTC
TCCTTGTCTCTACCCCTTCTCTGCTTTTCTGGGGCAGGGGCAAGTACCCCTCAACCCCTTCTCCTTCACCCTTAGCGG
TAAGTCCCGCTTTTCTAGGGGGCAGGGGCAAGTACCCCCCAACCCCTTCTCCTTCACCCTTAGTGGCAAGTCCTGCTT
TCCTGGGGCAGGGGCAAGTACCCCTTAACCCCTTCTCCTTCACCCTTAGCAGCAAGTCCCACTTTTCTAGGGGGCAAG
AAACCCCAAACCCCTTCCCTCCGTGTCTTTACGCTCTCTTTTCTCTGGGTTTGCTTCCTTCACTATGGGCAACCTTCC
ATCCTCCATTCCTCCTTCTCCCTTAGCCTGTGTGCTCAAGAACTTAAAACCTCTTCAACTCACACCTGACCTAAAACC
TAAATGCCTCATTTTCTTCTGCAACACCGCTTGGCCCCAATACAAACTTGACAATGGCTCTAAATGGCCAGAAAATGG
CACTTCGATTTCTCCATCCTACAAGACCTAAATAATTTTTGTCAAAAAATGGGCAAATGGTCTGAGGTGCCTGATGTC
CAGGCATTCTTTTACACATCCGTCCCTTCCTAGTCTCTGTGCCCAGTGCAACTCGTCCCAAATCTTCCTTCTTTCCCT
CCCACCTGTCCCCTCAGTCCCAACCCCAGGCGTTGCTGAGTGTGTCTAATCTTCCTTTTCTACAAACCCATCTGACCT
CTCCCCTCCTCTCCACGCCAAGCTAGGTCCCAATTCTTCCTCAGGCTCCACTCCTCCACCCTGTAATCTTTTTATCAC
CTCCCCTCCTCACACCTGGTCCGGCTTACAGTTTCGTTCCGTGACTAGCCCTCCCCCACCTGCCCAGCAATTTACTCT
TAAAAAGGTGGCTGGAGCTAAAGACATAGTCAAGGTTAATGCTCCTTTTTCTTTATCCCAAATCAGAAGCGTTTAGGC
TCTTTTTCATCAAATATAAAAACCCAGCCCAGTTCATGGCTCGTTCGGCAGCAACCCTGAGACGCTTTACAGCCCTAG
ACCCTAAAAGGTCAAAAGGCCGTCTTATTCTCAATAACATTTTATTACCCAATCTGCTCCCGATATTAAATAAAACTC
CAAAAATTAAATTCCAGCCCTCAAACTCCACAACAGGATTTAATTAACCTCGCCTTCAAGGTGTACAATAATaaaaaa
aaaaaaGTTGCAATTCCTTGCCTCCACTGTGAGACAAACCCCAGCCACATCTCCAGCAAACAAGAACTTCCAAACACC
TGAACCGCAGCAGCCAGGCGTTCCTCCAGAACCTCCTCCCACAGGAGCTTGCTACAAGTGCCAGAAATCTGGCCACCA
AGCCAAGAAATGCCTGCAGCCCAGGATTCCTCCTAAGCCGTGTCCCATCTGTGCGGGACCCCATTGGAAATCAGACTG
TTCAACTCACCTGGCAGCGACTCCCAGAGACCCTGGAACTCTGGCCCAAGGCTCTCTGACTGACTCCTTCCCAGATCT
TCTCGGCTTAGCGGCTGAAGACTGACGCTGCCCAATCACCTCGGAAGCCCTGTAGACCATCACAGACACTGAGCTTTA
GGTAACTCTCACAGTGGAGGGTAAGTCCGTCCCCTTCTTAATCAATATGGAGGCTCCCCACTCCACATTACCTACTTT
TCAAGGGCCTGTTTCCCTTGCCTCCATAACTGCTGTGGGTATTGACGGCCAGGCTTCTAAACCTCTTAAAACTTCCTC
TATCGACTGCTTACACAATACACATTCACGCACTCATTCCTAGTTATCCCCCCCTGCCCAGTTCTTATTAGGCCGAGA
CACTTTAACTAAATTATCTGCTTCCCTGACTACTCCTGGACTACAGCTGCATCTCATTGCCGCCCTTATTCCTAATCC
AAAGCCTCCTTTGCGTCCTCCTCTTGTATCCCCCCACCTTAACCCACAAGTATAAGATACCTCTACTCCCTCCTTGGC
GACCGATCATGCACCCCTTACCATCTCATTAAAACCTAATCACCCTTACCCCGCTCAATGCCAATATCTCATCCCACA
GCATGCTTTGAAAGGATTAAAGCCTGTTATCACTTGCCTGCTACAGCATGGGCTCCTAAAACCTATAAACTCTCCTTA
CCATTCCCCCATTTTACCTGTCCTAAAACCAGACAAGCCTTACAAGTTAGTTCAGAATCTGCGCCTTATCAACCAAAT
GTTTTGCCTATCCACCCCGTGGTGCCAAACCCGTATACTCTTCTATCCTCAATACCTCCCTCTACTACCCATTATTCT
GTTCTGGATCTCAAACATGCTTTCTTTACTATTCCTTTGCACCCTTCATCCCAGCCTCTCTTTGCCTTCACTTAGACT
GACCCTGACACCCATTAGGCTCAGCAAATTACCTGGGCTGTACTGCCGCAAGGCTTCACAGACAGCCCCCATTACTTC
AGTCAAGCCCAAATTTCATCCTCATCTGTTACCTATCTCAGCATAATTCTCATAAAAACACAGGTGCTCTCCCTGCTG
ATCGTGTCCGATTAATCTCCCAAACCTCAATCCCTTACAAAAGAACAACTCCTTTCCTTCCTAGGCATGGTTAGTGTG
GTCAGAATTCTTACACAAGAGCCAGGACTGCACCCTGTAGCCTTTCTGTCCAAACAACTTGACCTTCCTGTTTTAGCC
TAGCCATCATGTCTGTGTGCAGCCGCTGCCGCTGCTTTAATACTGTTAGAGGCCCTCAAAATCACAAACTATGCTCAA
CTCACTCTCTACATTTCTCATAACTTCCAAAATCTATTTTCTTCCTCATACCTGACGCATATACTTTCTGCTCCCCGG
CTCCTTCAGCTGTACTCACACTTTCTTAAGTCCCACAATTACCATTGTTCCTGGCCTGGACTTCAATCTGGCCTCCCA
CATTATTCCTGATACCACACCTGACCTCCATGACTGTATCTCTCTGATCCACCTGATATTCACCCCATATTTCCTTCT
TTCCTGTTCCTCACCCTGATCACGCTTGATTTATTGATGGAAGTTCCACCAGGCCTAATCGCCACACACCAGCAAAGG
CAGGTTATGCTATAGTACAAGCCACTAGCCCACCTCTCAGAACCTCTCATTTCCTTTCCATCGTGGAAATCTATCCTC
AAGGAAATAACTTCTCAGTGTTCCATCTGCTATTCTACTACTCCTCAGGGATTATTCAGGCCCCCTCCCTTCCCTACA
CATCAGGCTGGAGGATTTGCCCCACCCAGGACTGGCAAATTACCTTTACTCAACATGCCCTGATTCAGGAAACTAAAA
TACCTCTTAGTCTAAATAGACACTTTCACTGAATAAAGTAAAGGCCTTTCCTACAGGGTCTGAGAAGGCCACCACAGT
CATTTCTTCCCTTCTGTCAGACACAATTCCTCAGTTTAGCCTTCCCACCTCTATACAGTCTGATAACAGACCAGCCTT
TATTAGTCAAATCAGCCAAGCAGTTTTTCAGGCTCTTAGTATTCAGTGAAACCTTTATATCCCTTACGGTCCTCCTTC
TTCAAGAAAAGTAGAACGGACTAAAGGTCTTTTAAAAACACACCTCACCAAGCTCAGCCACCAGCTTAAAAAGGACTG
GACAATACTTTTACCACTTTCCCTTCTCAGAAATCAGGCCTGTCCTCAGAATGCTACAAGGTACATCCCATTTAAGCT
CCTGTATAGACGCTCCTTTTTATTAGGCCCCAGTCTCATTCCAGACACCAGACCAACTTAGACTGTGCCCCAAAAANC
NTGTCATCCCTACTATCTTCTGTCTAGTCATACTCCTATTCACCGTTCTCAACTACTCATACATGCCCTGCTCTTGTT
TACACTGCCGGTTTACACTGTTTCTCCAAGCCATCACAGCTGATATCTCCTGGTGCTATCCCCAAACTGCCACTCTTA
ACTCTTGAAGTAAATAATCTTTGCTGGCAGGACTATGCTGAATCTCCTTAGGCACTCTCTAATCAGATGTCCTAGGTC
CTCCCAATTCTTAGACATTTAATACCCATTTTTCTCCTCCTTTTATTCGGACCTTGTATCTTCCATTTAGCTTCTCAA
ATCATCCAATACCGNATCCAGGCCATCACCAATCCTTCTATACGANANATGTTTCTTCTAACATCCCCATGATATCAC
CCCTTACCACAAGACCTCCCTTCANCTTAATCTCTCCCACTCTAGGCTCCCACGCCGCCCCTAATCCCGCTTGAAGCA
GCCCTGAGAAACATCGCCCATTATCTCTCCATACCACCCCCCAAAAATTTTCGCTGCCCCAACACTTCAACACTATTT
TGTTTTATTTGTCTTATTAATATAAGAAGGCAGGAATGTCAGGCCTCTGAGCCCAGGCCAGGCCATCGCATCCCCTGT
GACTTGCACGTATACATCCAGATGGCCTAAAGTAACTGAAGATCCACAAAAGAAGTAAAAACAGCCTTAACTGATGAC
ATTCCAACATTGTGATTTGTTCCTGCCCCACCCTAACTGATAAATGTACTTTGTAATCTCCCCCACCCTTAAGAAGGT
CCTTTGTAATTCTCCCCACCCTTGAGAGTGTACTTNGTGAGATCCACACCTGCCCACCAGAGAACAAACCCCCTTTGA
CTGTAATTTTCCNTTACCTTCCCT

(B)

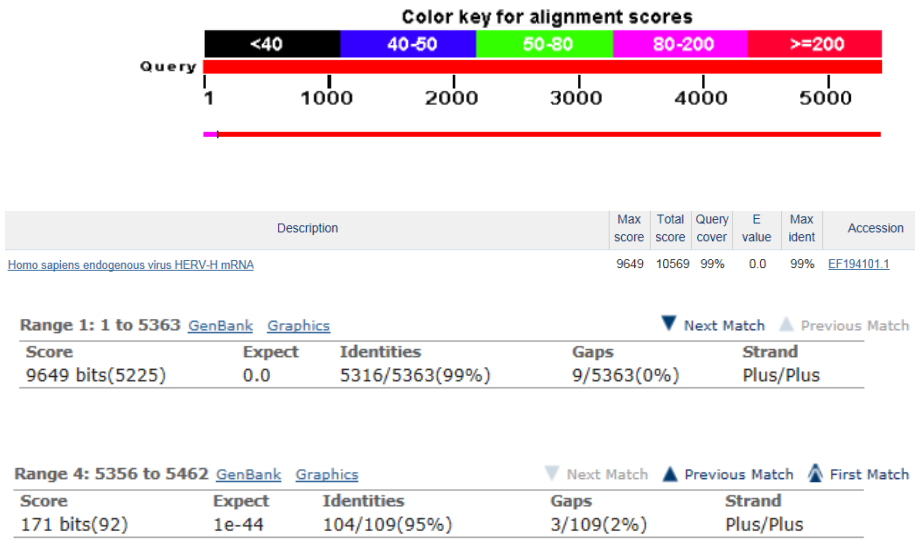
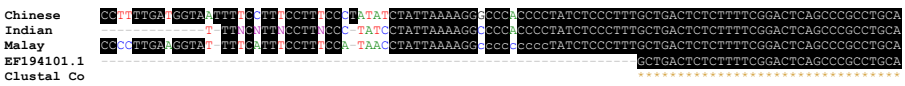
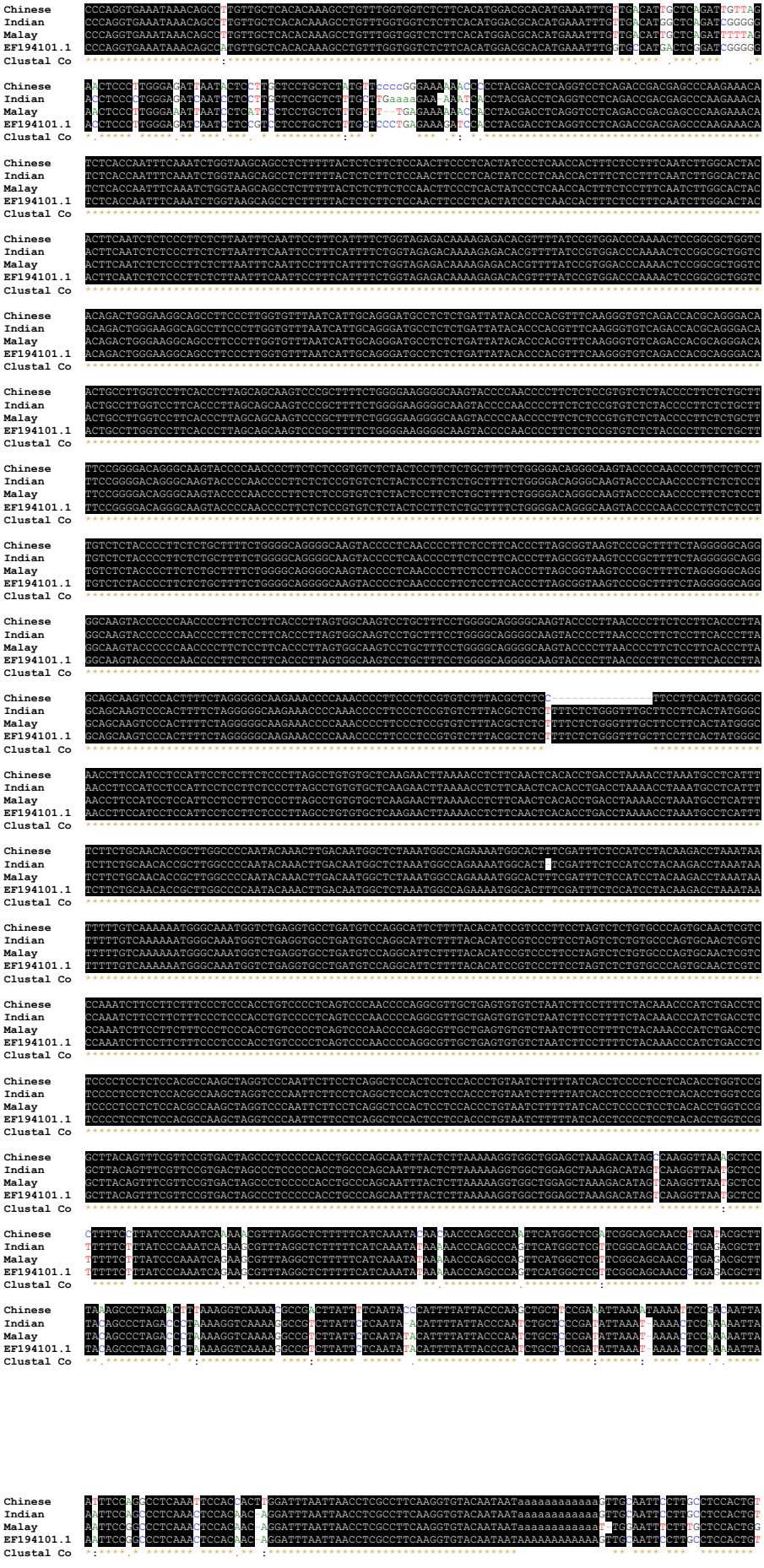


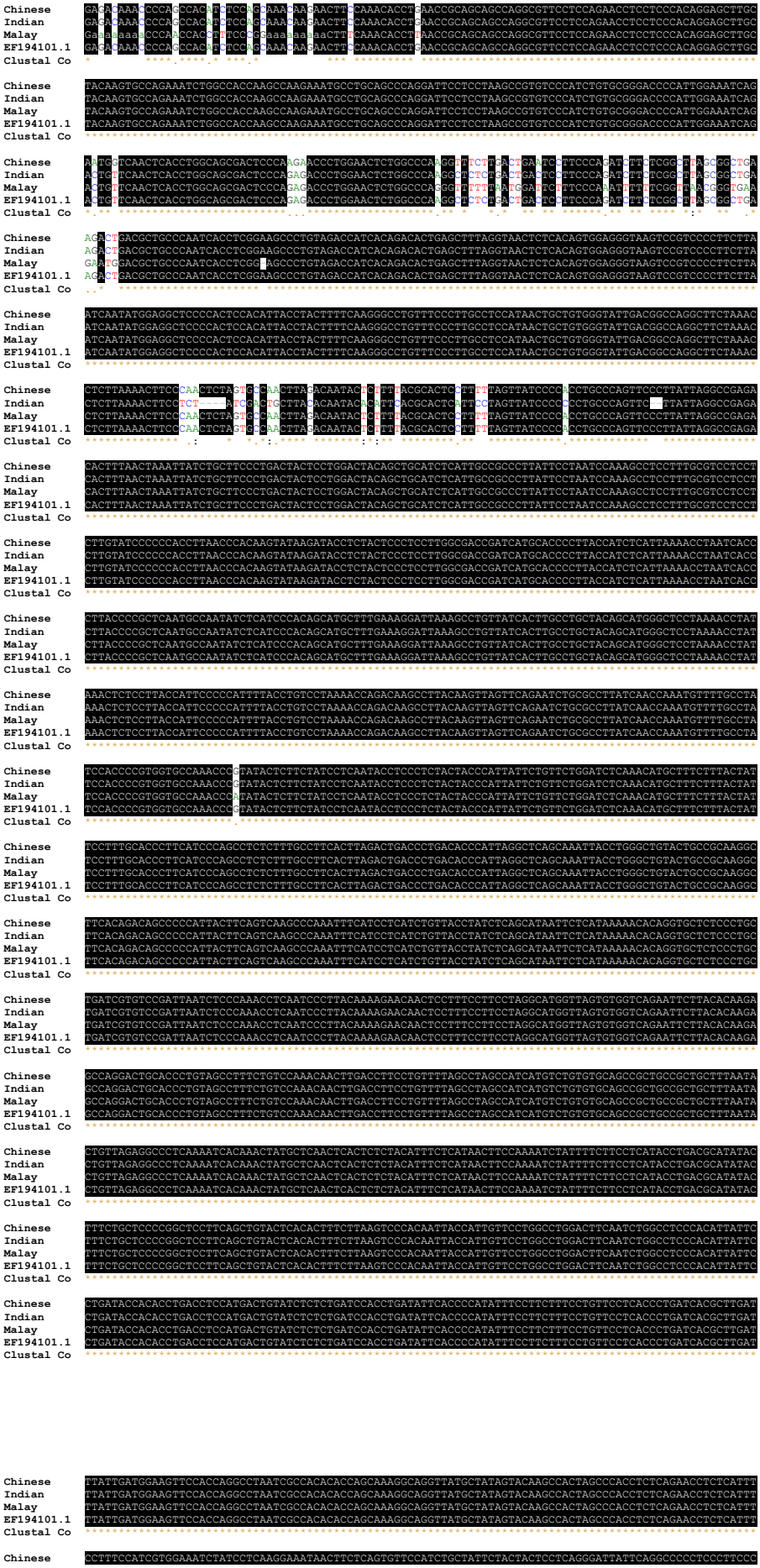
Figure 2.7. Definition of a HERV-H sequence from an Indian subject. (A) Assembly of a contiguous sequence of 6 kb HERV-H (B) Alignment scores from NCBI BLAST.

Table 2.4. Pairwise observed sequence identities (%). HERV-H sequence from GenBank: EF194101.1 and each of the three ethnic groups were compared with one another. Analysed data obtained from Clustal 2.1.

Observed identities (%)	Chinese	Indian	Malay	EF194101.1
Chinese	100	97.87	97.96	98.60
Indian	97.87	100	98.00	99.03
Malay	97.96	98.00	100	98.74
EF194101.1	98.60	99.03	98.74	100




Chinese
Indian
Malay
EF194101.1
Clustal Co


Chinese
Indian
Malay
EF194101.1
Clustal Co

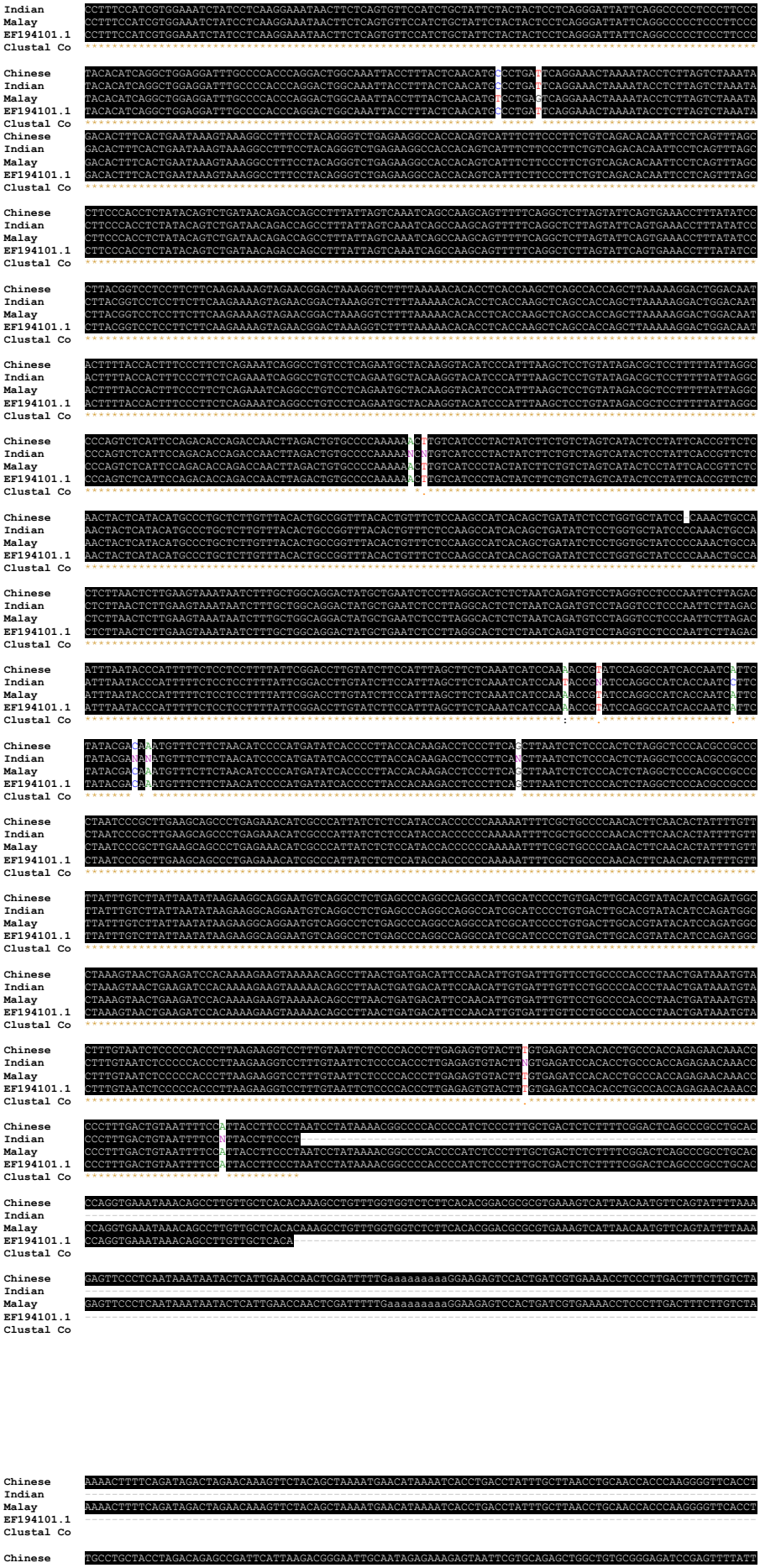


Figure 2.8. Comparison of HERV-H sequence. Result of pairwise alignment of HERV-H sequences from three different ethnic groups, namely, Chinese, Indian and Malay together with GenBank HERV-H (NCBI: EF194101.1) are shown. Clustal consensus sequence (Clustal Co) is also shown. Black background indicates sequence identity and white background indicates sequence dissimilarity.

2.5 Discussion

HERVs represent the trail of successful ancient retroviral infection of the human germ line. As a consequence of the persistence of vertical transmission, about 8% of the human genome is occupied by HERVs (Lander et al., 2001). Their existence, through millions years of evolution, implies that HERVs have overcome the stringent evolutionary selection and successfully domesticated into the host genomes although the accumulation of mutations, deletions, frame shifts or premature stop codons as well as hypermethylation of the promoter sites may have rendered them defective or inactive (Blikstad et al., 2008). While retroelements like Alu (in SINE) has been known to link to breast cancer and X-linked agammaglobulinemia, and L1 (in LINE) to haemophilia A and B, and Duchenne muscular dystrophy, as well as SVA to Fukuyama muscular dystrophy and neutral lipid storage disease with myopathy (Solyom and Kazazian, 2012, Cardelli and Marchegiani, 2012), the role of HERVs contributing to carcinogenesis remains unclear and perplexing.

To the author's knowledge, this is the first study presenting the prevalence of HERV-H in a Southeast Asian population. HERV-H was present in 89.3% of the sample Singapore population. Of late, one study involving 20 Western Asian subjects has reported that up to 87% of HERV-H insertional polymorphism was observed in all their subjects (Guliyev et al., 2013). These findings corroborate with the report that HERV-H, a Class I ERVs that is considered one of the oldest classes of ERVs (Class I and III), is occupying a relatively higher proportion in the human genome (Hohn et al., 2013).

Nevertheless, the differential prevalence rates of HERV-K elements in different human races (Woo et al., 2013, Jha et al., 2009, Turner et al., 2001) imply that the domestication process of retroviruses may not necessarily lead to the fixation of proviral DNA in the human genome. Although there may be times when the replicative activity of retrovirus is high and this can lead to the outbreaks of integration (Belshaw et al., 2005), the processes of evolution are capable of modulating genetic diversity. To elaborate, both selective forces and random genetic drift are capable of either fixing or losing the retroviral elements. In addition, population bottlenecks may also contribute to a lesser extent of genetic variation. It is noteworthy to find that "jumping genes" like retrotransposons are subjected to fixation or extinction in the human population during the course of evolution (Figure 2.3). Akin to gene knockout experiment *in vitro*, the insertional polymorphisms, fixed or loss of

endogenous retroviral elements in the genome, might provide the means of examining the role of HERV in diseases (Moyes et al., 2007).

A recent debate about the presence of HERV-K found in the genome sequences of archaic Neanderthal and Denisovan fossils (both ~40,000 years old), are indeed found in the genomes of modern humans reveals the process of integration is not completely fixed. To elaborate, HERV-K elements that were found in Neanderthals and Denisovans, were initially reported to be absent in modern human reference genome. With more analyses performed on many new genome sequences of modern-day humans, it is realised that these HERV-K elements are indeed present in these modern-day humans. This demonstrates that the HERV-K elements are present in the genome at fluctuating frequencies, a result of the effects of genetic drift over the years (Marchi et al., 2013, Agoni et al., 2012). On the same note, the loss of HERV-H gene could occur over many generations as evidenced in human evolution (Mager and Freeman, 1995). Taken together, the high prevalence rate of HERV-H in the sample Singapore population provides a logical basis to conduct further studies on the role of HERV-H in colorectal carcinogenesis. To put it in another way, a low distribution frequency or even extinction of HERV-H in the population genetics of Singapore would sensibly negate the needs to further investigate the causal role of HERV-H in oncogenesis.

Knowing that HERV-H is highly prevalent in the sample local population, a further close examination of the three contiguous sequences from 3 different ethnic groups was performed and revealed a high level of

conservation, up to 99%. Generally, highly conserved sequences are regarded as functional and biologically important. However, it is unclear why these proviral sequences remain undisturbed and well-fixed over long periods of evolutionary history. A thorough evaluation of the biological significance of this evolutionarily conserved region is beyond the current scope of this study, but this observation obviously warrants further investigations. On this note, putative coding regions were analysed and expression of such a putative peptide has led to some interesting findings with regard to HERV-H in carcinogenesis. These findings will be discussed in the following chapter.

In summary, there is a high prevalence of HERV-H in the sample Singapore population and this high distribution frequency is independent of factors like gender. Additionally, the percentage of sequence identities for HERV-H between different ethnic groups is generally very high, suggesting that the HERV-H region is well conserved. The findings of this study could therefore be the basis for future HERV-H molecular and population genetic studies.

Role of HERV-H in colorectal carcinogenesis

Chapter 3

CHAPTER 3

ROLE OF HERV-H IN COLORECTAL CARCINOGENESIS

3.1 Introduction

Cancer is a devastating disease and incurs an immerse toll on those afflicted everywhere (Bray et al., 2013). It is a leading cause of death in most countries and the escalating healthcare expenses associated with cancer treatment (Mariotto et al., 2011, Warren et al., 2008) exact huge financial burdens on patients and their families. The substantial amounts of money spent on cancer research (Eckhouse et al., 2008) has not abated as healthcare administrators continue to support the search for new drugs and better treatment options for cancer patients. In Singapore, cancer was among the leading causes of death in 2006 – 2011 (Teo and Soo, 2013). Hence, the burden of cancer presents a major public health problem that requires continuous governmental support to combat the scourge of cancer, especially when the multifactorial pathogenesis of carcinogenesis is always evolving (Hanahan and Weinberg, 2011).

3.2 Colorectal cancer

According to GLOBOCAN 2008, a database created under the auspices of the International Agency for Research on Cancer, World Health Organization, colorectal cancer, cancer of the large bowel (colon and rectum), is the third most common cancer in men and the second in women (Ferlay et

al., 2010). Approximately 60% of the cases occur in developed countries. Incidence rates vary considerably in both sexes worldwide, the highest rates being in Australia/New Zealand and Western Europe, the lowest in Africa (except Southern Africa) and South-Central Asia, and intermediate in Latin America. Incidence rates are substantially higher in men than in women (overall ratio 1.4:1). It accounts for 8% of all cancer deaths, making it the fourth most common cause of death from cancer.

Over the past four decades, Singapore has shown a dramatic increase in the incidence rate of colorectal cancer. With an annual increasing average rate of approximately 2.6% and 2.35% for men and women respectively, colorectal cancer is now the most frequent cancer when both genders are combined (Wong and Eu, 2007, Ministry of Health, 2010, Teo and Soo, 2013).

The development of colorectal cancer is associated with many contributing risk factors. An overview of these risk factors is summarised in Table 3.1. The classic description of colorectal carcinogenesis is the adenoma-carcinoma sequence and multistep tumorigenesis, together with various genetic alterations and several biological pathways (Fearon and Vogelstein, 1990, Markowitz and Bertagnolli, 2009, Cappell, 2008). An overview of these hereditary genetic alterations is represented in Table 3.2. It is noteworthy that the majority of colorectal cancers (up to 85%) do not develop in association with hereditary genetic variations. In fact, these sporadic colorectal cancers involves the accumulation of sequential mutations of several key signalling pathways, mediating the transitional phase of normal mucosa-to-benign

adenoma-to-severe dysplasia-to-frank carcinoma. *APC* mutations, *K-ras* mutations, loss of 18q (including *DCC*, *SMAD2* and *SMAD4*) and deletion of 17p (*TP53*), together with DNA methylation and malfunction of the mismatch repair genes, are the molecular events associated with sporadic colorectal cancers (Cunningham et al., 2010).

Table 3.1. Risk factors for colorectal cancer (Cappell, 2008, Centers for Disease Control and Prevention, 2011, American Cancer Society, 2012)

Parameters	Proposed mechanism
Epidemiology	
Old age	Acquired colonocyte mutations accumulate with age
Living in highly industrialised nations, e.g. US	Dietary and environmental carcinogens
Diet	
Low fruit and vegetable	Anticarcinogenic substances in fruits and vegetables (e.g., folic acid)
Obesity	Carcinogens in an unhealthy diet or possible role of abnormal insulin levels in carcinogenesis
Social habits	
Smoking cigarettes	Carcinogens present in tobacco; associated with MSI
Alcohol	May promote cell proliferation, inhibit DNA repair, contribute abnormal DNA methylation; suppress tumour immune surveillance, alter composition of bile acids, induce cytochrome p450 enzymes to activate carcinogens or low levels of folic acid
Genetics/family history	
FAP (familial adenomatous polyposis)	Develops hundreds of adenomatous colonic polyps. Inevitably develops colon cancer resulting from small but significant risk for malignant transformation in each adenoma
Gardner's syndrome	Variant of FAP
HNPCC (Lynch syndrome)	Mutant mismatch repair gene leads to accumulation of genetic mutations, including mutations of tumour suppressor genes
Peutz-Jeghers syndrome	Syndromic hamartomatous polyps occasionally may transform to adenomas
Juvenile polyposis	Syndromic juvenile polyps can transform to adenomas and then cancers over time
Family history of nonsyndromic colon cancer	Postulated shared genetic factors leading to mild susceptibility to colon cancer and possibly shared environmental factors
Hyperplastic polyposis	Genetic mutation in hyperplastic polyposis seems to predispose to colon cancer

Inflammatory bowel disease	
Chronic ulcerative colitis	Dysplasia and genetic mutations associated with mucosal injury and repair
Chronic Crohn's colitis	Dysplasia and genetic mutations associated with cell injury and repair
History of prior neoplasia	
Colonic adenomatous polyps	Precursor lesions of colon cancer
Prior colon cancer	Genetic predisposition or environmental factors
Other	
Pelvic radiation	Carcinogenic effects resulting from radiation-induced mutations
Streptococcus bovis bacteraemia	May promote colonocyte proliferation
Ureterosigmoidostomy	Carcinogens excreted in urine or colonic mucosal proliferation during repair after urine-induced mucosal injury
Acromegaly	Growth hormone promotes proliferation of preexisting colonic adenomas and cancers
Diabetes mellitus	Insulin may modulate colonocyte proliferation
Prior cholecystectomy	Continuous colonic exposure to potentially carcinogenic bile acids after cholecystectomy
Lifestyle factors	
Physical inactivity	Physical activity may stimulate immunosurveillance and stimulate intestinal peristalsis to decrease mucosal contact with faecal carcinogens
Low calcium	Calcium binds to bile acids that otherwise are potentially colonotoxic
High fat	Various theories (e.g., increased bile secretion to induce cell proliferation)
High red meat	Animal fat in red meat or carcinogens (e.g., nitrosamines)
Low selenium	Selenium can help neutralize toxic free radicals due to antioxidant effects
Low folate	Folate needed for DNA synthesis and repair
Low carotenoid diet	Carotenoids can help neutralize free radicals resulting from antioxidant effects
Low fiber diet	Dilution of carcinogens in stool due to increased stool bulk and stool water with a high fibre diet

Table 3.2. Inherited colorectal cancer syndromes (Cheah, 2009, Gryfe, 2009)

Syndrome	Associated genes
Hereditary non-polyposis colorectal cancer (HNPCC) or Lynch syndrome	*MSH2, MLH1, MSH3, MSH6, PSM1, PSM2*

Familial adenomatous polyposis (FAP) and Variants: - Gardner - Turcot - Attenuated FAP	*APC*
Hamartomatous polyposis syndromes	
Juvenile polyposis syndrome (JPS)	*SMAD4, BMPR1A*
Peutz-Jeghers syndrome (PJS)	*LKB1*
Cowden syndrome	*PTEN*
Hereditary mixed polyposis syndrome (HMPS)	*BMPR1A, CRAC1*
MUTYH-associated adenomatous polyposis (MAP)	*MYH*
Hyperplastic/serrated polyposis syndrome	Yet to be elucidated

3.3 Oncogenic viruses

Viral infections have been reported as aetiological agents in various diseases. In recent years, many researchers have put forward the notion that infectious agents are capable of transforming the normal cells, and directing the infected host to malignancy. Oncogenic viruses like Epstein–Barr virus (EBV), hepatitis B virus(HBV), human papillomaviruses (HPV), human T-cell lymphotropic virus type 1 (HTLV-1), hepatitis C virus (HCV), and Kaposi's sarcoma-associated herpesvirus (KSHV) are the 6 viruses classified by the International Agency for Research on Cancer as being "carcinogenic to humans". To this end, approximately 11% of cancers are attributable to exogenous viruses (Parkin et al., 1999).

Human endogenous retroviruses have been associated with the pathogenesis of cancer. Retrovirus-related elements have been reported to be detected in cancers like seminomas, testicular teratocarcinomas and choriocarcinomas as well as certain breast cancers, small-cell lung

carcinomas, renal cancers and leukaemias (Ryan, 2004). However, the role of these transcripts remains poorly defined. On the other hand, studies have indicated that immune responses to the human endogenous retroviral proteins are closely associated with the progression or the prognosis of cancer. These include the HERV-K *gag* protein in prostate cancer (Reis et al., 2013) and the HERV-K *env* protein in breast cancer (Zhao et al., 2011).

3.4 Human endogenous retroviruses

The human endogenous retroviruses were first reported in 1981 (Martin et al., 1981). Approximately 8% of the human genome is occupied by HERVs (Bannert and Kurth, 2004, Lander et al., 2001) and this was thought to be the result of retroviral proviruses inserting themselves into the germline of human ancestors 40 million years ago (Barbulescu et al., 1999). Obviously, the initial infection of the germline must have overcome several factors which include the host immune system (Baumann, 2006) that encompasses innate antiretroviral protein activity like TRIM 5α (Towers, 2005), APOBEC3G (Cullen, 2006) and siRNA (Blikstad et al., 2008). In the course of evolution, the accumulation of mutations, truncations, frame shifts or premature stop codons as well as hypermethylation of the promoter sites may have rendered the retroelements to be defective or inactive (Blikstad et al., 2008). However, there remains a potential that these retroelements may affect the stability of the genome through illegimate recombination or hypomethylation (Schulz et al., 2006).

While the studies of retroviruses have enhanced our understanding on the molecular biology of cancer, development, differentiation and gene regulation, research on endogenous retroviruses is just beginning. Notably, studies on the association of cancers with HERVs have focused on the expression of retroelements. The discovery of Rec and Np9 of HERV-K has shed some light on tumorigenesis. Both Np9 and the Rev-like regulatory protein, Rec, interact with the promyelocytic leukaemia zinc finger protein PLZF (Boese et al., 2000), resulting in an abrogation of transcriptional repression of the c-*myc* gene. As a consequence, increased cell proliferation and cell survival ensue.

3.4.1 HERV-H in colorectal cancer

Recent clinical findings reported that HERV-H is selectively expressed in colon cancers but not in normal tissues (Alves et al., 2008, Liang et al., 2009, Liang et al., 2007b, Wentzensen et al., 2004, Wentzensen et al., 2007). Specifically, different transcripts of HERV-H, namely, *gag* (Alves et al., 2008) and *env* (Wentzensen et al., 2004) were observed in these studies involving patient samples (Table 3.3). Concordance between the findings of these four clinical studies performed at different geographical locations with divergent populations involved strongly suggests that HERV-H may play a role in colorectal carcinogenesis. The findings from these studies also revealed HERV-H was expressed in at least 50% of colorectal cancers. In fact, the study by Liang and colleagues reported 95% of their tumour cases were associated with HERV-H expression (Liang et al., 2009). Additionally, about

22% of colorectal adenomas were found to demonstrate HERV-H transcriptional activity.

Table 3.3. Clinical findings of HERV-H expression in colorectal cancers.

Sample size	HERV-H up-regulation	Transcript	Reference
21 tumour-normal tissue matched pairs	16/21 (76%)	*env*	(Wentzensen et al., 2004)
14 tumour-normal tissue matched pairs	7/14 (50%) colorectal carcinoma	*gag*	(Wentzensen et al., 2007)
34 tumour-normal tissue matched pairs	18/34 (53%) colorectal carcinoma	*gag*	
36 tumour-normal tissue matched pairs	8/36 (22%) colorectal adenoma	*gag*	
25 tumour samples	11/18 (60%) primary colorectal carcinoma 7/8 (88%) metastases 1/1 (100%) pre-cancerous adenoma	*gag*	(Alves et al., 2008)
20 tumour-normal tissue matched pairs	19/20 (95%)	*gag*	(Liang et al., 2009)

Although the regulatory mechanism of the expression of HERV mRNA remains poorly understood, it is postulated that transducing agents like hormonal and environmental signals may activate the 5' long terminal repeat (LTR) of HERV and lead to consequential downstream transcriptional activities (Wentzensen et al., 2007). Based on the ORFs of HERV-H transcripts, a putative gag protein of 93 amino acids would be generated, but it is not known how this putative protein would contribute to tumour

progression. Interestingly, a recent finding has demonstrated that the expression of HERV-K transmembrane envelope proteins inhibits immune cell activation and modulates IL-10 release. This suggests that the transmembrane protein may suppress the immune system and thereby prevent the rejection of tumour cells and facilitate the process of carcinogenesis (Morozov et al., 2013).

Nevertheless, the role of HERV-H in colorectal carcinogenesis remains to be elucidated. To investigate how HERV-H is involved in tumorigenesis, we examined the effect of HERV-H overexpression on human colorectal cancer cell lines. Specifically, the ability to proliferate, migrate and invade was studied as these characteristics demonstrate important aspects of *in vivo* tumour biology. A knockdown assay was also conducted to study the effect of reduced HERV-H expression on colorectal cancer cells. In addition, the expression levels of CD133 and CD44, the widely-used cell surface markers for colon cancer stem cells (O'Brien et al., 2007, Dalerba et al., 2007), were also examined. These studies sought to characterise the mechanisms by which HERV-H contributes to colorectal carcinogenesis.

3.5 Materials and Methods

3.5.1 Cell lines and culture conditions

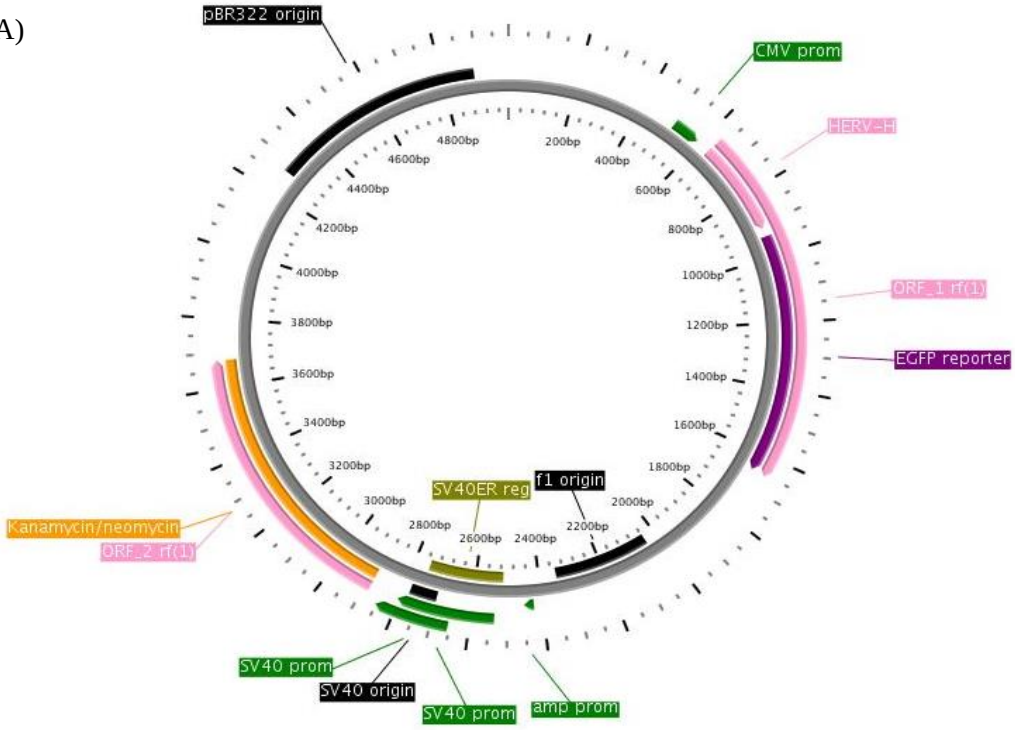
HT29, a human colon adenocarcinoma cell line from a 44-year-old Caucasian female (Fogh, 1975) and LS174T, a human colon adenocarcinoma cell line from a 58-year-old Caucasian female (Tom et al., 1976) were obtained from the NUH-NUS Tissue Repository at the National University

Hospital. HT29 cells and LS174T cells were routinely cultured in McCoy's 5A Medium supplemented with 10% fetal bovine serum (FBS) and in Minimum Essential Medium (MEM) α supplemented with 10% fetal bovine serum, respectively. Both cell lines were maintained in a humidified atmosphere of 5% CO_2 in air at 37˚C and used when in the log phase of growth.

3.5.2 Plasmid generation

The plasmid pEGFP-N2 (Clontech, Palo Alto, California, United States) was a kind gift from Dr Wang Cheng, Duke-NUS Graduate Medical School Singapore. pEGFP-N2:HERV-H was constructed by amplification of *HERV-H* with the primer pair A (5′- ctc AAG CTT atg ggc aac ctt cca tcc t -3′) and B (5′- cgt GGA TCC cgg aag gga cgg atg tgt aa -3′) to provide HindIII and BamHI restriction sites without the stop codon, respectively. The restriction sites for the primer pair are represented by uppercase letters. The amplified fragment was digested with HindIII and BamHI and inserted into the HindIII and BamHI sites of the pEGFP-N2 vector. The inserted fragment of the construct was sequenced and verified. The resulting pEGFP-N2:HERV-H encoded a 93 amino acid-HERV-H tagged with a green fluorescent protein (GFP). Figure 3.1 illustrates the vector pEGFP-N2:HERV-H.

(A)



(B)

Atgggcaacc ttccatcctc cattcctcct tctcccttag cctgtgtgct caagaacttа
aaacctcttc aactcacacc tgacctaaaa cctaaatgcc tcatttтctt ctgcaacacc
gcttggcccc aatacaaact tgacaatggc tctaaatggc cagaaaatgg cactttcgat
ttctccatcc tacaagacct aaataatttt tgtcaaaaaa tgggcaaatg gtctgaggtg
cctgatgtcc aggcattctt ttacacatcc gtcccttccG ggatccaccg gccggtcgcc
accatggtga gcaagggcga ggagctgttc accggggtgg tgcccatcct ggtcgagctg
gacggcgacg taaacggcca caagttcagc gtgtccggcg agggcgaggg cgatgccacc
tacggcaagc tgaccctgaa gttcatctgc accaccggca agctgcccgt gccctggccc
accctcgtga ccaccctgac ctacggcgtg cagtgcttca gccgctaccc cgaccacatg
aagcagcacg acttcttcaa gtccgccatg cccgaaggct acgtccagga gcgcaccatc
ttcttcaagg acgacggcaa ctacaagacc cgcgccgagg tgaagttcga gggcgacacc
ctggtgaacc gcatcgagct gaagggcatc gacttcaagg aggacggcaa catcctgggg
cacaagctgg agtacaacta caacagccac aacgtctata tcatggccga caagcagaag
aacggcatca aggtgaactt caagatccgc cacaacatcg aggacggcag cgtgcagctc
gccgaccact accagcagaa cacccccatc ggcgacggcc ccgtgctgct gcccgacaac
cactacctga gcacccagtc cgccctgagc aaagacccca acgagaagcg cgatcacatg
gtcctgctgg agttcgtgac cgccgccggg atcactctcg gcatggacga gctgtacaag
taa

(C)

Sequence ID: lcl|51309 Length: 1023 Number of Matches: 1

Range 1: 1 to 1023 Graphics

Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
1890 bits(1023)	0.0	1023/1023(100%)	0/1023(0%)	Plus/Plus

```
Query  14    ATGGGCAACCTTCCATCCTCCATTCCTCCTTCTCCCTTAGCCTGTGTGCTCAAGAACTTA  73
Sbjct  1     ATGGGCAACCTTCCATCCTCCATTCCTCCTTCTCCCTTAGCCTGTGTGCTCAAGAACTTA  60

Query  74    AAACCTCTTCAACTCACACCTGACCTAAAACCTAAATGCCTCATTTTCTTCTGCAACACC  133
Sbjct  61    AAACCTCTTCAACTCACACCTGACCTAAAACCTAAATGCCTCATTTTCTTCTGCAACACC  120

Query  134   GCTTGGCCCCAATACAAACTTGACAATGGCTCTAAATGGCCAGAAAATGGCACTTTCGAT  193
Sbjct  121   GCTTGGCCCCAATACAAACTTGACAATGGCTCTAAATGGCCAGAAAATGGCACTTTCGAT  180

Query  194   TTCTCCATCCTACAAGACCTAAATAATTTTTGTCAAAAAATGGGCAAATGGTCTGAGGTG  253
Sbjct  181   TTCTCCATCCTACAAGACCTAAATAATTTTTGTCAAAAAATGGGCAAATGGTCTGAGGTG  240

Query  254   CCTGATGTCCAGGCATTCTTTTACACATCCGTCCCTTCCGGGATCCACCGGCCGGTCGCC  313
Sbjct  241   CCTGATGTCCAGGCATTCTTTTACACATCCGTCCCTTCCGGGATCCACCGGCCGGTCGCC  300

Query  314   ACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTG  373
Sbjct  301   ACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTG  360

Query  374   GACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACC  433
Sbjct  361   GACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACC  420

Query  434   TACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCC  493
Sbjct  421   TACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCC  480

Query  494   ACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATG  553
Sbjct  481   ACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATG  540

Query  554   AAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATC  613
Sbjct  541   AAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATC  600

Query  614   TTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACC  673
Sbjct  601   TTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACC  660

Query  674   CTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGG  733
Sbjct  661   CTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGG  720

Query  734   CACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAG  793
Sbjct  721   CACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAG  780

Query  794   AACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTC  853
Sbjct  781   AACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTC  840

Query  854   GCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAAC  913
Sbjct  841   GCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAAC  900

Query  914   CACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATG  973
Sbjct  901   CACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATG  960

Query  974   GTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAG  1033
Sbjct  961   GTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAG  1020

Query  1034  TAA  1036
Sbjct  1021  TAA  1023
```

Figure 3.1. The pEGFP-N2:HERV-H expression construct. (A) Schematic drawing of pEGFP-N2:HERV-H plasmid map. The amplified 298bp-HERV-H is inserted into the multiple cloning site of the pEGFP-N2 vector. With CMV early promoter, HERV-H is synthesized as a fusion protein, tagging on to the N terminus of EGFP. A neomycin resistance cassette (Neo^r), consisting of the SV40 early promoter, the kanamycin/neomycin resistance gene of *Tn5*, and polyadenylation signals from the Herpes simplex virus thymidine kinase gene, allows stably transfected eukaryotic cells to be selected using G418. Plasmid map was generated using PlasMapper Version 2.0 (Dong et al., 2004) (B) The sequence of fusion HERV-H:GFP in plasmid. Blue highlight: Start/stop codon; Grey highlight: HERV-H; Green highlight: GFP. (C) NCBI Nucleotide BLAST analysis of actual sequencing result of fusion HERV-H:GFP (http://blast.ncbi.nlm.nih.gov/Blast.cgi) is shown.

3.5.3 Transfection and stably transfected cell lines generation

Transfection of HT29 and LS174T cells was carried out in serum-free McCoy's 5A and MEMα respectively, using Lipofectamine 2000 CD (Invitrogen, USA) in a 6-well plate according to the manufacturer's instructions. Stably transfected HT29 and LS174T cells were generated by selecting the successfully integrated *neo* gene using the antibiotic G418 at the concentration of 800 μg/mL. Selection took place in a 6-well plate, where it took approximately 3-4 weeks for foci of G418-resistant colonies to appear. Thereafter, the resistant foci were clonally expanded over an additional 2–3 weeks. Colonies were then picked and expanded on a 6-well plate, followed by growth in a T25 tissue culture flask.

After 3–4 weeks, adherent transfected cells were trypsinized into single-cell suspensions and sorted by flow cytometry (DB FACSVantage SE™). To facilitate the recovery of large numbers of transfected cells, pre-sorting triggered by fluorescence was carried out. Using forward scatter as a triggering signal, the final sort was grouped into GFP-positive and GFP-negative populations. The resulting GFP positive cells were resuspended in complete medium with G418 and cultured in a humidified atmosphere of 5% CO_2 in air at 37˚C and used when in the log phase of growth.

3.5.4 Subcellular localization

For subcellular localization assays, transfected cells were grown to about 30% density on glass coverslips and fix in 1% paraformaldehyde in

phosphate-buffered saline (PBS) for 20 minutes at room temperature. Following 1X PBS wash, the cells were stained with Hoechst 33342 at a final concentration of 5μg/mL for 15 minutes in the dark. Stained cells were imaged under the Olympus IX71 fluorescent microscope, equipped with cellSens Standard 1.5 software (Olympus).

3.5.5 Proliferation assay

Cell proliferation assays were carried out by seeding cells at a density of 2×10^5 viable cells to each well of a 6-well plate. At the indicated time points, the cells were trypsinized and counted using a Luna™ automated cell counter (Logos Biosystems, Korea) after trypan blue staining. All results were presented as means ± standard deviation from triplicate wells.

3.5.6 Invasion assay

Invasion assays were performed in 24-well plate transwell chambers with 8μm-pore polycarbonate filter inserts (the CHEMICON cell invasion kit). Briefly, a total of 1.25 x 10^5 cells were seeded on the ECM-coated inserts in 50 μL of serum-free medium inserts for invasion assays. The lower chambers were filled with 500 μL of 10% FBS-supplemented McCoy's 5A or MEMα medium. After 24 hours, cells on the upper side of the filter were removed by aspiration and the cells on the lower surface of the insert were dislodged by cell detachment solution. Detached cells were subsequently lysed and stained with CyQuant GR Dye. An aliquot of the lysis/dye mixture was

transferred to a 96-well plate for fluorometric reading using 485/535 nm filter set. The RFU value represents the invasive ability. Assays were performed in triplicates. Data were presented as mean RFU values ± standard deviation from triplicate wells.

3.5.7 Scratch assay

HT29 cells transfected with HERV-H:GFP or with GFP (2×10^5) were seeded in triplicate in a 6-well tissue culture plate and cultured in McCoy's 5A containing 10% fetal bovine serum, to form confluent cell monlayers. A single scratch was then made with a 10μL pipette tip across the cell layer and rinsed twice with PBS to remove non-adherent cells. The wounded monolayers were subsequently cultured in medium containing 10% fetal bovine serum at 37°C. The repopulation of the denuded area was monitored by capturing images every 24 hours for 72 hours at fixed positions. All cell cultures were analysed in parallel and triplicate measurements were taken. Cell migration was assayed by evaluating closure of a linear wound over these 72 hours. Areas were quantified by image analysis using Wimasis image analysis software (ibidi GmbH, Germany). The percentage of wound (scratch area) at 0 h (control) was arbitrarily assigned as 100%.

3.5.8 Sphere-forming ability of LS174:HERV-H_GFP

Adherent colorectal cells were trypsinized into single-cell suspensions and were plated into 6-well plates with a density of 2×10^5 cells per well.

Cells were grown and complete medium was refreshed every 3 days. Sphere formation was monitored at 3, 6, 9 and 12 days. At day 12, pictures were taken with an inverted microscope (IX70 Olympus). The diameter of the spheres in all wells was measured using DP2-BSW software (Olympus) from saved images captured at 10x magnification. An analysis of large sphere (>100 μm) formation ability was analysed.

3.5.9 Revival of HERV-H_GFP- and GFP-transfected LS174T and HT29 cells after 30 days of serum deprivation

Adherent colorectal cells were trypsinized into single-cell suspensions and were plated onto 6-well plates with a density of 2 x 10^5 cells per well. Cells were allowed to attach for 24 hours in medium supplemented with 10% fetal bovine serum. Thereafter, cells were grown in medium without fetal bovine serum. Medium was refreshed at every 5-day intervals for 25 days. At day 30, the remaining adherent cells in the 6-well plates were refreshed with medium supplemented with 10% fetal bovine serum. This day was designated as post-famine day 0. Cell growth was monitored at 0, 2, 4 or 6 days post-famine and the colonies were photographed in two randomly chosen microscopic fields per well, using the IX70 inverted microscope (Olympus). The diameter of the colonies in all wells was measured using DP2-BSW software (Olympus) from saved images captured at 10x magnification. Experiments were done in triplicates.

3.5.10 Detection of CD44 and CD133 expression by flow cytometry

Adherent colorectal cells were trypsinized into single-cell suspensions and rinsed once with PBS. Dissociated cells were stained with PE-conjugated anti-CD44 antibody or anti-CD133 antibody (Miltenyi Biotech, USA) and incubated for 30 minutes in the dark at 4°C. Mouse IgG1-PE (Miltenyi Biotech, USA) was the isotype control antibody. After incubation, the cells were washed once in PBS and resuspended in 500 μL cold PBS with 10% FBS for flow cytometry analysis within 1 hour. The flow cytometry analysis of cells was carried out using the FACSCalibur flow cytometry system (Decton Dickinson, USA) with 50,000 events being counted for each case.

3.5.11 Knockdown of HERV-H by Dicer-Substrate small interfering RNA (DsiRNA)

3.5.11.1 DsiRNA Design

A total of three DsiRNA targeted against HERV-H mRNA were designed (Integrated DNA Technology, Coralville, IA, USA). The sequences of the three DsiRNAs are as follows:

(1) WH 1:

Sense sequence 5’ – CCU CCU UCU CCC UUA GCC UGU GUG C – 3’

Antisense sequence 5’ –GCA CAC AGG CUA AGG GAG AAG GAG GAA– 3’

(2) WH 2:

Sense sequence 5’ – CUC CAU CCU ACA AGA CCU AAA UAA T– 3’

Antisense sequence 5’ – AUU AUU UAG GUC UUG UAG GAU GGA GAA– 3’

(3) WH 3:

Sense sequence 5’ – CAA AUG GUC UGA GGU GCC UGA UGT C– 3’

Antisense sequence 5’ – GAC AUC AGG CAC CUC AGA CCA UUU GCC– 3’.

A validated negative control duplex (scrambled DNA) with the following sequences was used:

Sense sequence 5’ –CGU UAA UCG CGU AUA AUA CGC GUA T– 3’

Antisense sequence 5’ –AUA CGC GUA UUA UAC GCG AUU AAC GAC– 3’

3.5.11.2 DsiRNA Transfection

The day before transfection, cells were trypsinized, diluted with fresh medium and transferred to 6-well plates (2 x 10^5cells/well). 2nM or 10nM DsiRNAs were transfected using transfection reagent (Lipofectamine RNAiMax) according to manufacturer’s protocol (Invitrogen, USA). Routinely, transfection was carried out at about 10–20% cell confluency as

recommended.

3.5.11.3 Effect of HERV-H targeting DsiRNA at the mRNA level

At 72 hours after DsiRNA transfection, cells were harvested and total RNA was prepared by RNeasy® mini kit (Qiagen, Germany). Harvested total RNA was quantitated using Nanodrop 100 spectrophotometer (ThermoScientific, USA) and 1μg of the total RNA was then reverse transcript using the iScript™ cDNA Synthesis Kit (Bio-Rad Laboratories, USA) in accordance to the manufacturer's instructions. Briefly, 1μg of the extracted RNA was mixed with enzyme reverse transcriptase and buffer to a final volume of 20μL and subjected to thermal profile of 25ºC for 5 minutes, 42ºC for 30 minutes followed by 85ºC for 5 minutes, in accordance to the manufacturer's instructions.

The primers used for the house-keeping gene β actin were the forward primer: 5′- ACC AAC TGG GAC GAC ATG GAG AAA-3′ and the reverse primer 5′-TAG CAC AGC CTG GAT AGC AAC GTA-3′.

The primers used for HERV-H were the forward primer: 5'- CTT CCC TCC GTG TCT TTA CG-3' and the reverse primer: 5'- AAG ATT AGA CAC ACT CAG CAA CG-3'.

The quantitative real time polymerase chain reaction (qRT-PCR) was performed using the iTaq™ Universal SYBR® Green Supermix (Bio-Rad Laboratories, USA) on the BioRad CFX96^{TM} Real-Time PCR system (Bio-Rad Laboratories, USA). Briefly, 1μL of cDNA and 0.5μL of the forward and

the reverse primers were added to iTaq™ Universal SYBR® Green Supermix. The reaction mix was then subjected to thermal profile of denaturation at 95ºC for 30 seconds, followed by amplification and quantification in 40 cycles at 95ºC for 5 seconds followed by 60ºC for 30 seconds. At the end of amplification cycles, melting temperature analysis was performed by the BioRad CFX96TM Real-Time PCR system (Bio-Rad Laboratories, USA). Relative gene expression was quantified based on $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001).

3.5.11.4 Effect of DsiRNA constructs on cell proliferation

At 72 hours after DsiRNA transfection, the effect of WH1, WH2 and WH3 DsiRNA on cell viability was determined by trypan blue assay using a Luna™ automated cell counter (Logos Biosystems, Korea).

3.5.12 Statistical methods

Values are expressed as mean plus or minus standard deviation (SD) for 3 experiments. Unless specified, data differences were subject to statistical analysis by paired student's t-test using STATGRAPHICS® Centurion XVI version 16.1.05 (Statpoint Technologies, USA). *P* values of <0.05 were considered statistically significant. All statistical tests were based on 2-sided probability.

3.6 Results

3.6.1 Subcellular localization of HERV-H

To examine the subcellular localization of HERV-H, a fusion construct of enhanced green fluorescent protein (GFP) linked to the N terminus of 93-amino-acid HERV-H was generated. The colocalization of fusion construct HERV-H:GFP was studied and it was found that HERV-H:GFP fusion protein demonstrated cytoplasmic staining (Figure 3.3). Previous studies have also shown that retroviral gag proteins are predominantly in the cytosol (Yu et al., 1995, Resh, 2005). For the control, free GFP was observed to be expressed throughout the cell including nuclei (Figure 3.2).

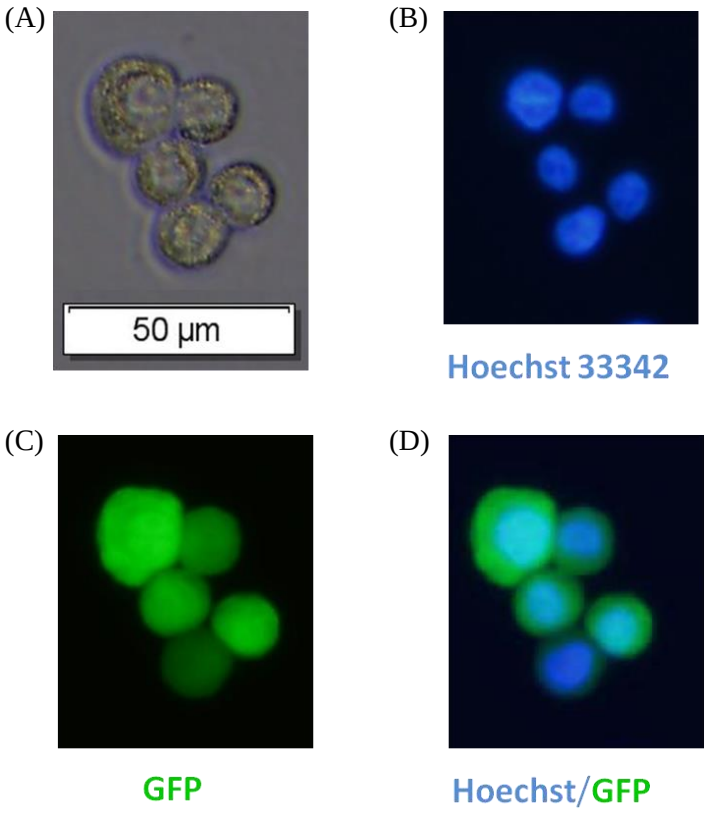


Figure 3.2. Intracellular localization of GFP. HT29 cells were transfected with plasmid pEGFP-N2 producing the indicated free GFP. (A) Phase contrast image of transfected HT29. (B) Nuclei of transfected HT29 were stained with Hoechst 33342. (C) Transfected HT29 analyzed for GFP fluorescence. (D) The merged image shows colocalization of nuclei and GFP.

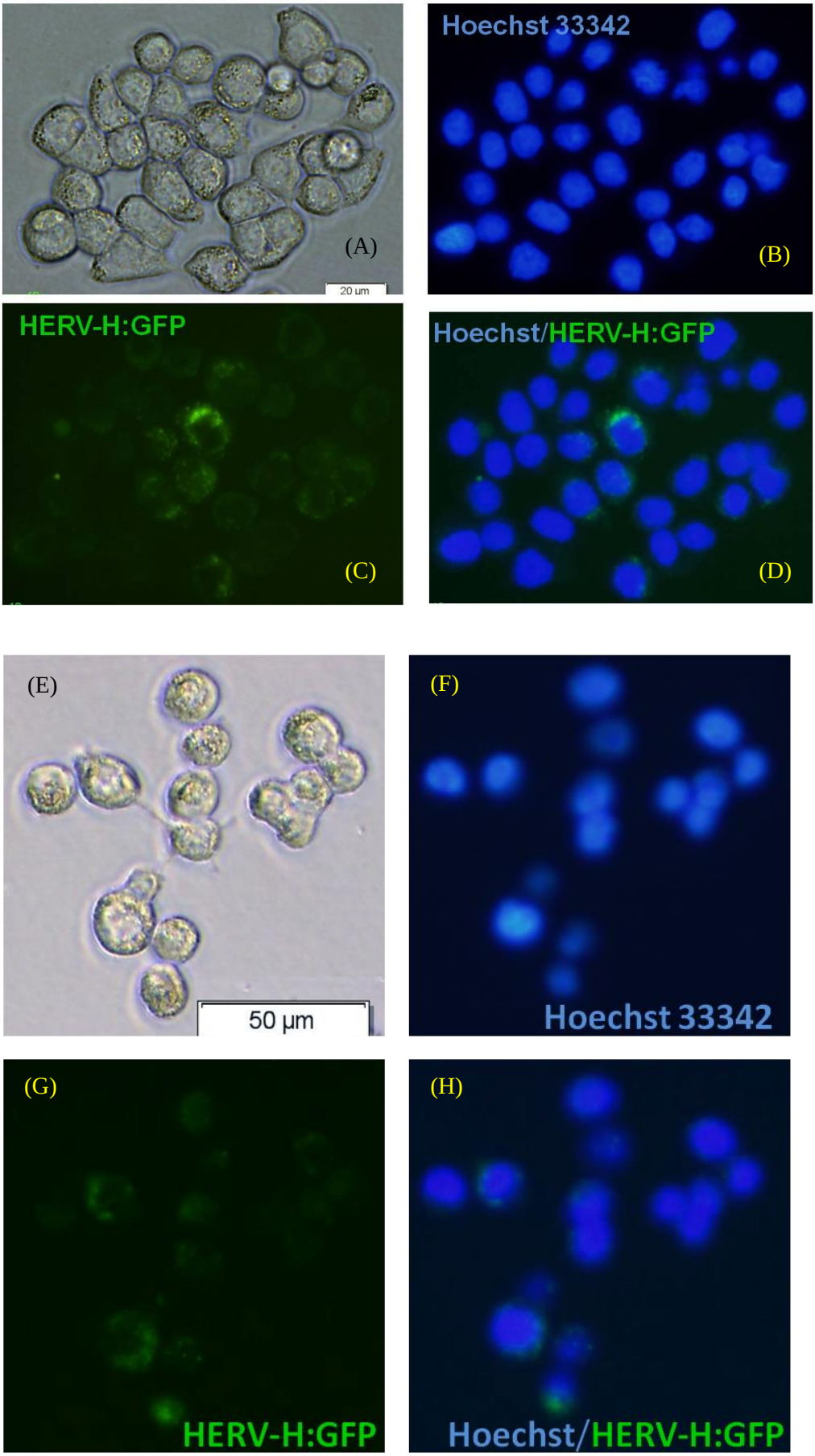
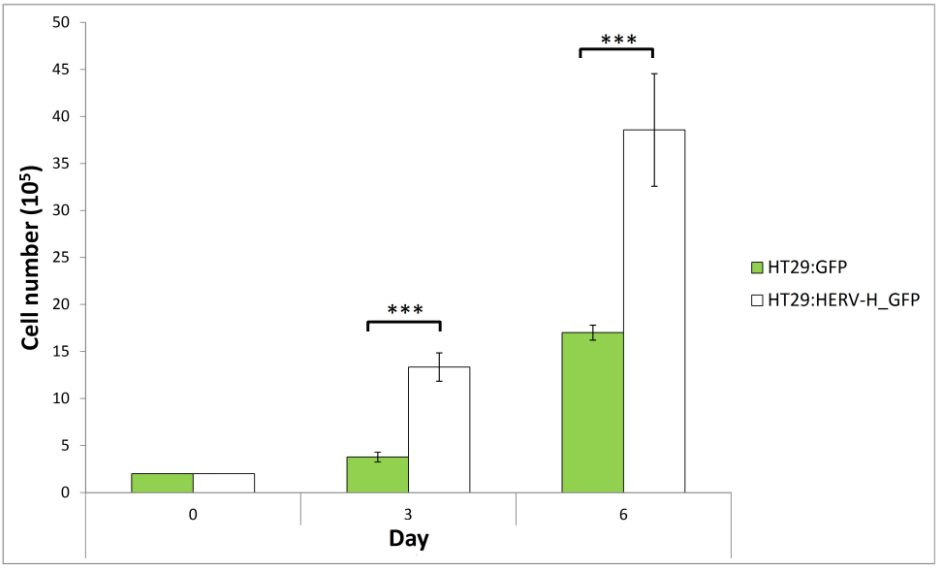

(A)
20 um
Hoechst 33342
(B)
HERV-H:GFP
(C)
Hoechst/HERV-H:GFP
(D)
(E)
50 µm
(F)
Hoechst 33342
(G)
HERV-H:GFP
(H)
Hoechst/HERV-H:GFP

Figure 3.3. Intracellular localization of HERV-H. HT29 cells were transfected with plasmid pEGFP:HERV-H producing the fusion protein HERV-H tagged with GFP. Cytoplasmic HERV-H:GFP is stable in transfected cells. (A,E) Phase contrast image of transfected HT29. (B,F) Nuclei of transfected HT29 were stained with Hoechst 33342 (C,G) Transfected HT29 analyzed for HERV-H:GFP fusion protein fluorescence. (D,H) The merged image shows colocalization of nuclei and HERV-H:GFP fusion protein.

3.6.2 Proliferation assay

To investigate the effect of HERV-H overexpression on the rate of proliferation, HT29 and LS174T cells were transfected with pEGFP-N2:HERV-H to upregulate the HERV-H expression levels. Compared to the control group which is the respective cell line transfected with empty plasmid pEGFP-N2, HERV-H overexpression enhances the proliferative potency of colorectal cancer cells significantly (Figure 3.4, $P<0.05$ and $P<0.001$).

(A)



(B)

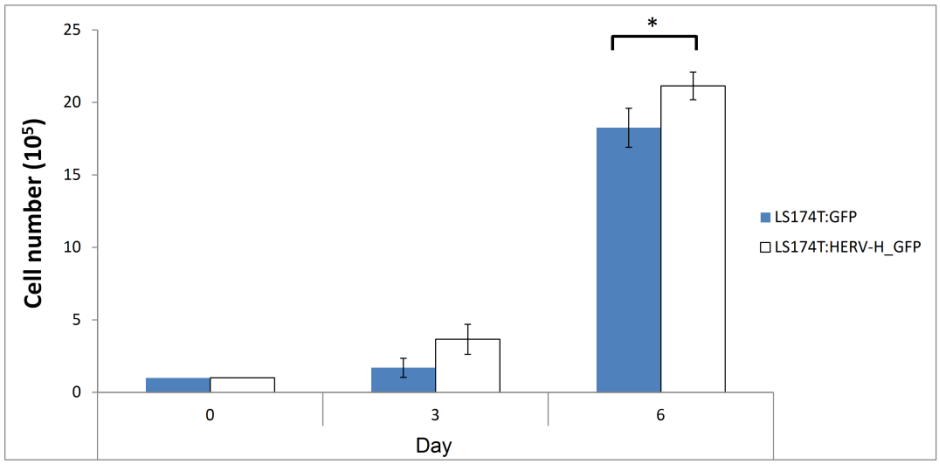
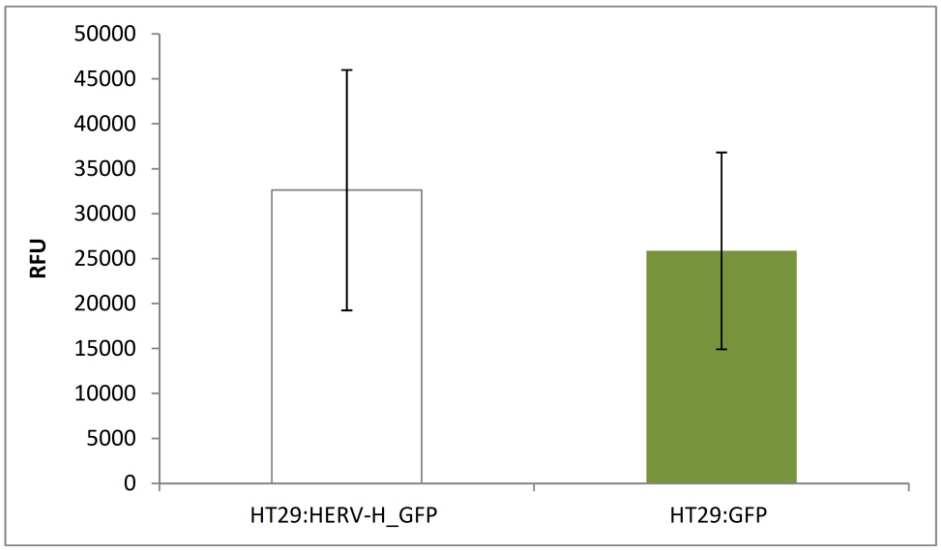


Figure 3.4. Effect of HERV-H overexpression on the rate of proliferation of HT29 and LS174T cells. (A) HT29 cells were transfected with plasmid pEGFP-N2:HERV-H or empty plasmid pEGFP-N2. (B) LS174T cells were transfected with plasmid pEGFP-N2:HERV-H or empty plasmid pEGFP-N2. At the indicated time points, the cells were trypsinized and counted using a Luna™ automated cell counter (Logos Biosystems, Korea) after Trypan blue staining. Experiments were done in triplicates. ***, P <0.001; *, P < 0.05.

3.6.3 Invasion assay

Cell invasion through the extracellular matrix (ECM) is an important process during tumour metastasis. Tumour cells initiate invasion by adhering to and spreading along the blood vessel wall. The invasion capability of cells is revealed by their ability to penetrate into the ECM-coated filter membrane in the transwell invasion chamber assay. The invasion potential of the HERV-H transfected colorectal cancer cells was examined and the results obtained revealed there was no observable difference in the invasiveness potential of the HERV-H transfected HT29 and LS174T cells and the respective control cells which were transfected with empty plasmid pEGFP-N2 (Figure 3.5).

(A)



(B)

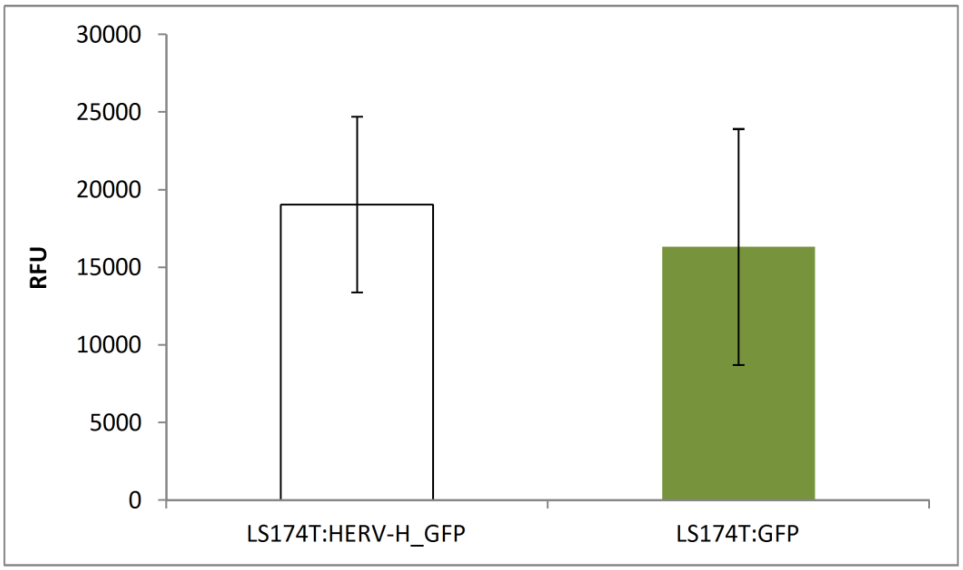


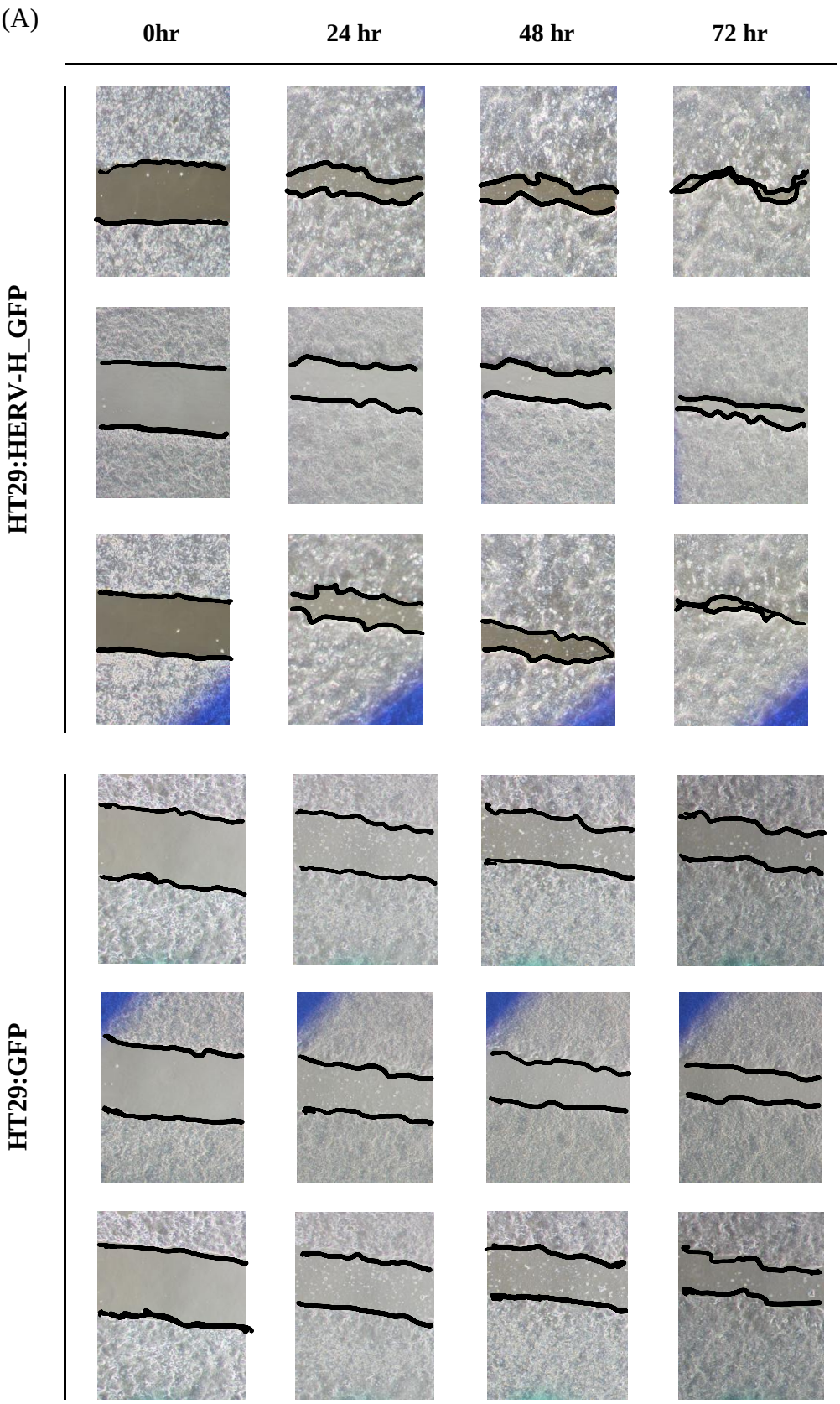
Figure 3.5. HERV-H does not promote colorectal cancer cell invasion. Cells were allowed to invade towards 10% FBS for 24 h. Fluorescence measurements were taken as described. Data were presented as mean RFU values ± standard deviation from triplicate wells. (A) HT29 cells were transfected with plasmid pEGFP-N2:HERV-H or empty plasmid pEGFP-N2. (B) LS174T cells were transfected with plasmid pEGFP-N2:HERV-H or empty plasmid pEGFP-N2. There is no significant difference in RFU values for both transfected cell lines.

3.6.4 Scratch assay

The contributory role of HERV-H in cell migration was examined using the *in vitro* scratch assay. The *in vitro* scratch assay is widely regarded as a straightforward method to study cell migration (Rodriguez et al., 2005, Lampugnani, 1999). With basic steps like the creation of a "scratch", a gap made on a confluent monolayer, and meticulous examination of the microphotographs taken during the closing of the scratch, the rate of cell migration can be measured to indicate cell motility (Liang et al., 2007a). The inherent cell morphology of LS174T cells is clumpy rather than uniform. Hence, it is unlikely to obtain a confluent monolayer of LS174T cells without gaps. In view of this technical limitation, only HT29 cells were subjected to scratch assay.

The percentage of migration area covered after 72 h was 78 ± 7% for HERV-H expressing HT29 cells and 40 ± 13% for control cells (Figure 3.6). The percentage of migration area of the HERV-H expressing HT29 cells was significantly higher than that of the control cells ($P < 0.05$). The results revealed the HERV-H plays a role in the cell migration properties of colorectal cancer cells.

(A)



(B)

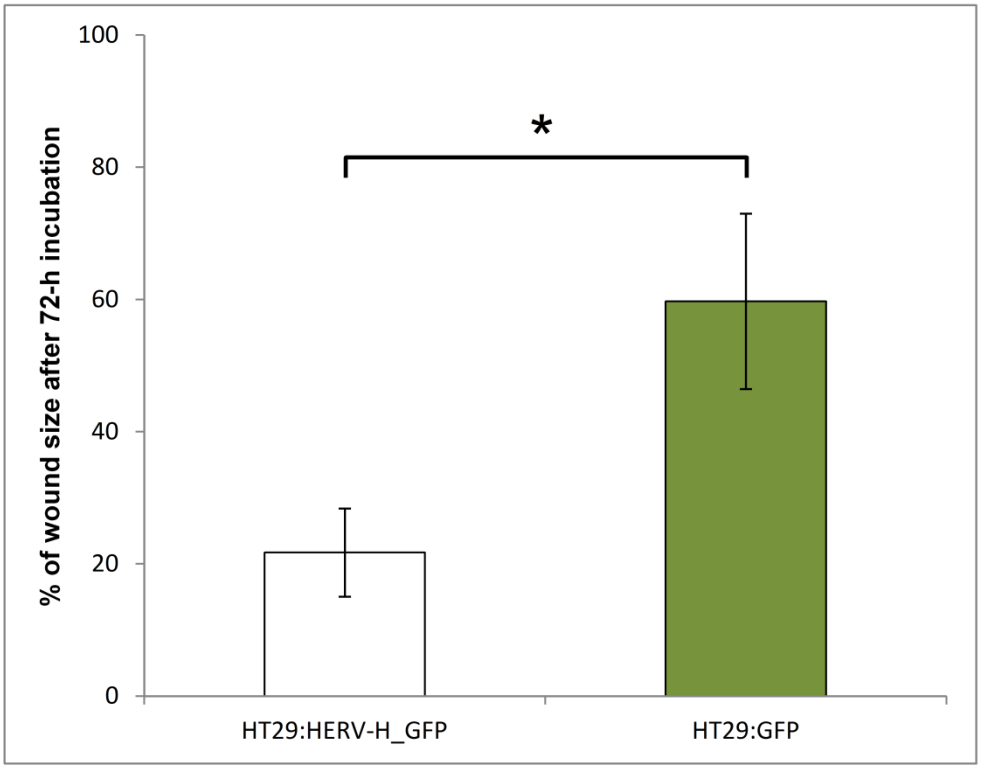


Figure 3.6. Cell migration of HT29 cells over a 72-h period in a wound scratch assay. (A) Representative images of cells migrating into the wounded area at 0, 24 h, 48 h and 72 h. (B) Graph showing the percentage of wound in control (HT29:GFP) and transfected cells (HT29:HERV-H_GFP) after 72-hour. The percentage of wound (scratch area) at 0 h (control) was arbitrarily assigned as 100%. Quantitative analysis was done using the WimScratch software (Wimasis GmbH, Munich, Germany). Results are presented as mean ± SD of three independent experiments. *P <0.05 when compared with control (HT29:GFP).

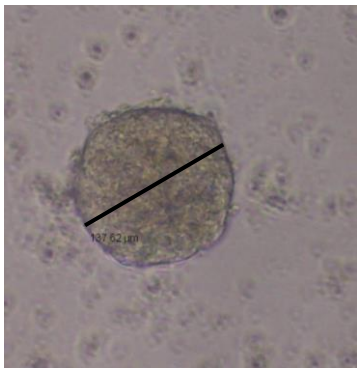
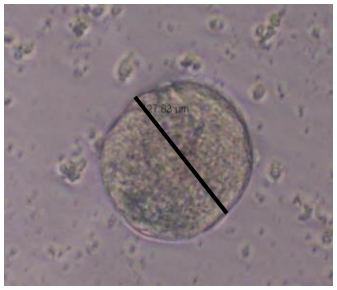
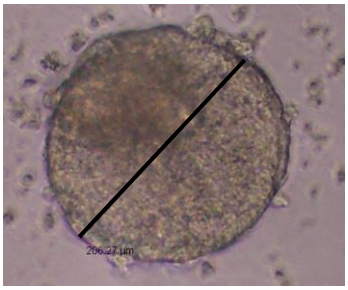
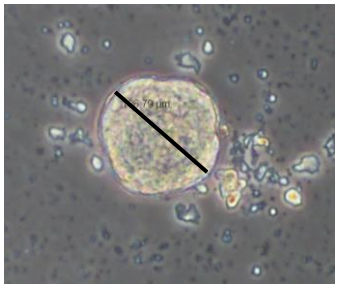
3.6.5 HERV-H induces large sphere forming ability in LS174T cells

The ability to form a sphere is an indicator of a cancer stem-like phenotype (Lobo et al., 2007). To examine the ability of HERV-H to induce sphere-forming ability in colorectal cells, HERV-H expressing LS174T cells were cultured and refreshed at 3-day interval. It was found that when compared to the control group, HERV-H expressing LS174T cells notably

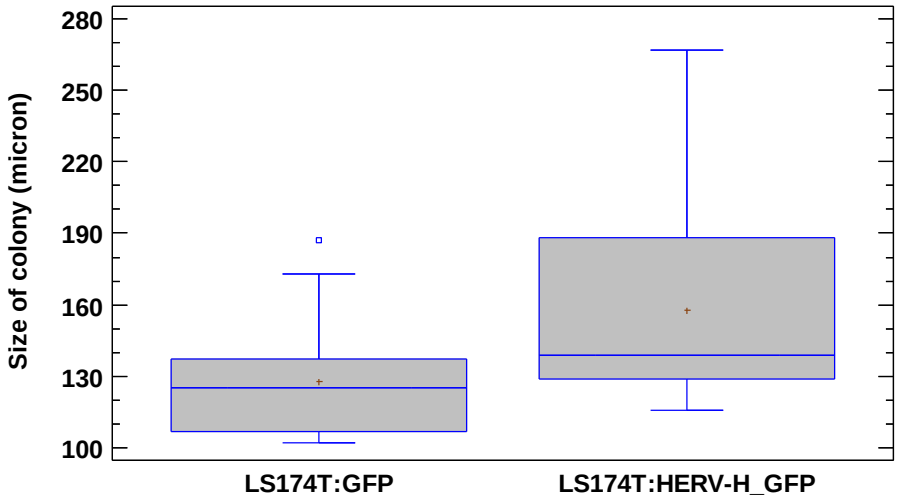
possess the ability to form large spheres (>100 μm). Using the Mann-Whitney-Wilcoxon test, the large-sphere forming ability of LS174T:HERV-H_GFP is significantly higher when compared to that of LS174T:GFP ($P < 0.01$) (Figure 3.7).

(A)

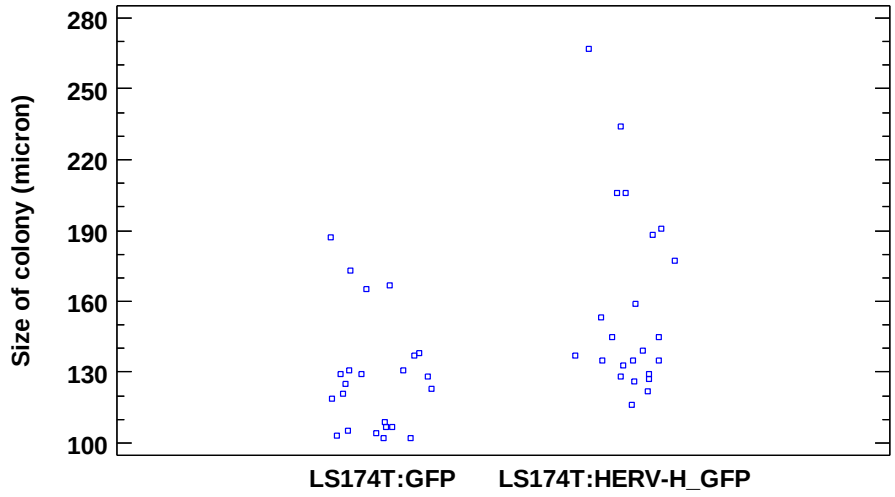
Spheres

LS174T:GFP	**LS174T:HERV-H_GFP**
	
∅137.62 μm	∅266.78 μm
	
∅127.83 μm	∅206.27 μm
	
∅106.79 μm	∅205.94 μm

(B)



(C)



Comparison of Medians
Median of sample 1: 125.0
Median of sample 2: 139.0

Mann-Whitney (Wilcoxon) W-test to compare medians
Null hypothesis: median1 = median2
Alt. hypothesis: median1 NE median2

Average rank of sample 1: 17.2174
Average rank of sample 2: 29.7826

W = 409.0 P-value = 0.00155082
Reject the null hypothesis for alpha = 0.05.

Figure 3.7. Sphere-forming ability of LS174T:HERV-H_GFP. (A) Representative images of large sphere (>100 μm) are shown. Actual diameter ∅ in μm is shown. Photomicrographs (X15 magnification) were taken by IX70 microscope (Olympus) equipped with DP2-BSW software (Olympus) (B) Box-and-Whisker plot showing 25th, 50th and 75th percentiles (horizontal lines) of spheres that are >100 μm. The plus sign (+) indicates the location of the sample mean. (C) Dot plot showing the spread of spheroid sizes. Mann-Whitney W-test indicates the ability of LS174T:HERV-H_GFP to form large sphere is significantly different ($P < 0.01$) from that of LS174T:GFP.

3.6.6 Revival of HERV-H_GFP- and GFP-transfected LS174T and HT29 cells after 30 days of serum deprivation

To investigate the effect of HERV-H overexpression on the dependence of serum for growth, the survival rate of HERV-H overexpressing LS174T and HT29 cells after 30 days of serum deprivation was tested (Figure 3.8). Interestingly, HERV-H overexpressing LS174T and HT29 cells survived and were able to form large colonies following serum replacement (Figures 3.9 and 3.11). Using the Mann-Whitney-Wilcoxon test, the colony-forming ability of LS174T:HERV-H_GFP is significantly higher when compared to that of LS174T:GFP ($P < 0.01$) (Figure 3.9). Similarly, the colony-forming ability of HT29:HERV-H_GFP is significantly higher when compared to that of HT29:GFP ($P < 0.05$) (Figure 3.11).

When compared to control cells, the growth rate of HERV-H overexpressing LS174T and HT29 cells was significantly higher after serum replacement ($P <0.001$) (Figures 3.10 and 3.12).

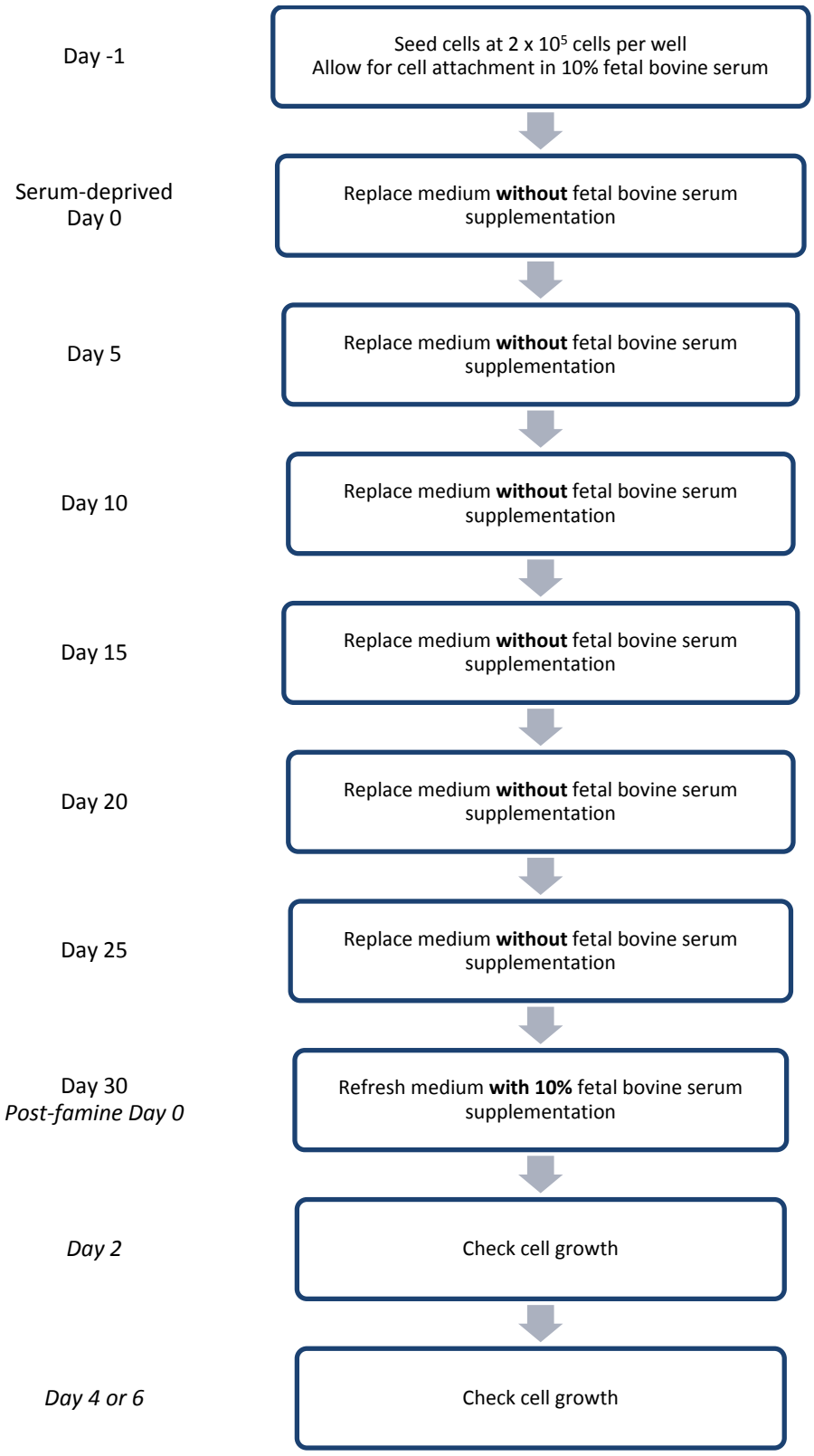
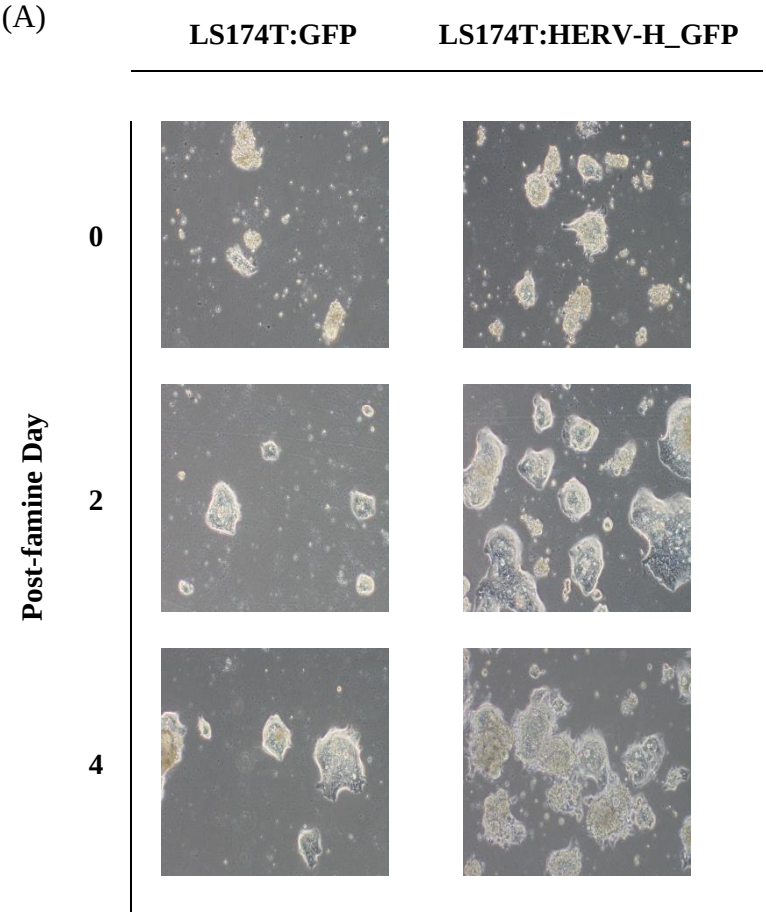
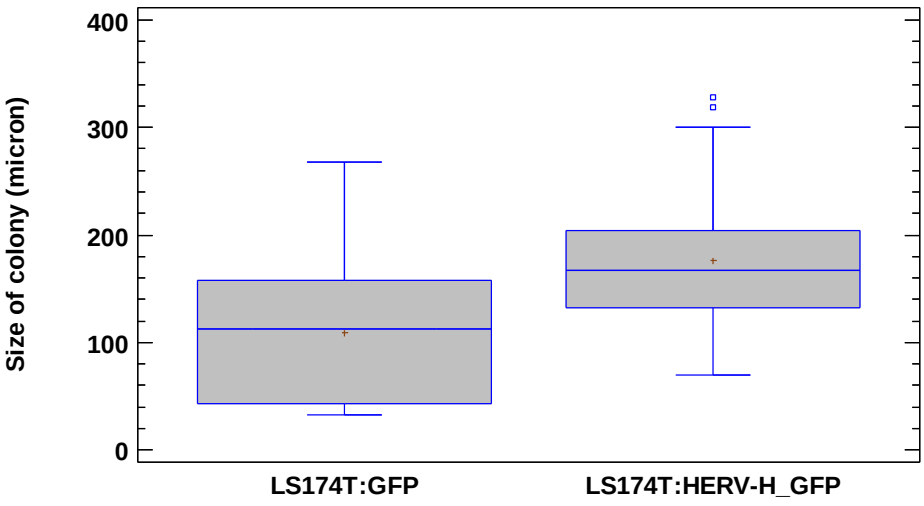


Figure 3.8. The timeline of revival assay for HERV-H transfected cell lines

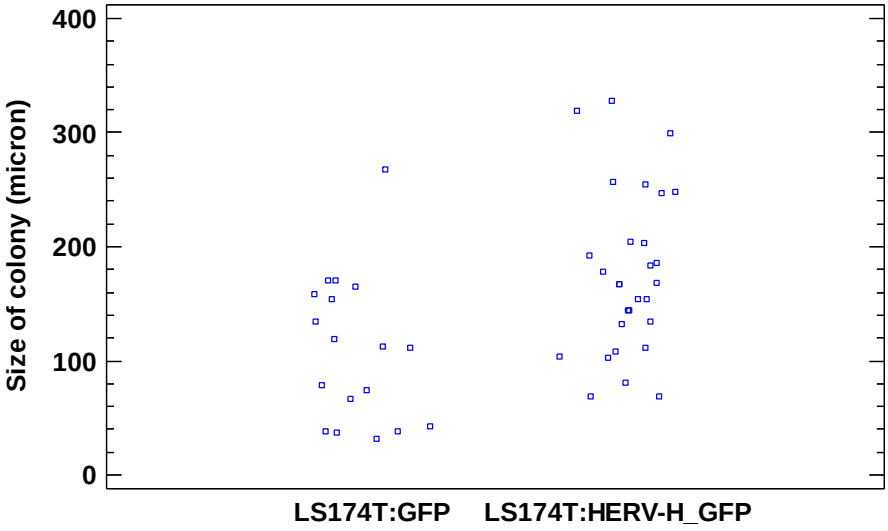
(A)


LS174T:GFP
LS174T:HERV-H_GFP
Post-famine Day
0
2
4

(B)


Size of colony (micron)
400
300
200
100
0
LS174T:GFP
LS174T:HERV-H_GFP

(C)



Comparison of Medians
Median of sample 1: 112.0
Median of sample 2: 167.0

Mann-Whitney (Wilcoxon) W-test to compare medians
Null hypothesis: median1 = median2
Alt. hypothesis: median1 NE median2

Average rank of sample 1: 16.6667
Average rank of sample 2: 28.5517

W = 393.0 P-value = 0.00399311
Reject the null hypothesis for alpha = 0.05.

Figure 3.9. Survival of transfected LS174T cells following 30-day of serum deprivation. (A) Representative images of respective transfected LS174T cells at indicated time points are shown. Photomicrographs (X10 magnification) were taken by IX70 microscope (Olympus) equipped with DP2-BSW software (Olympus). (B) Box-and-Whisker plot showing 25th, 50th and 75th percentiles (horizontal lines) of various sizes of colonies. The plus sign (+) indicates the location of the sample mean. (C) Dot plot showing the spread of colony sizes. Mann-Whitney W-test indicates the ability of LS174T:HERV-H_GFP to form numerous large colonies is significantly different ($P < 0.01$) from that of LS174T:GFP.

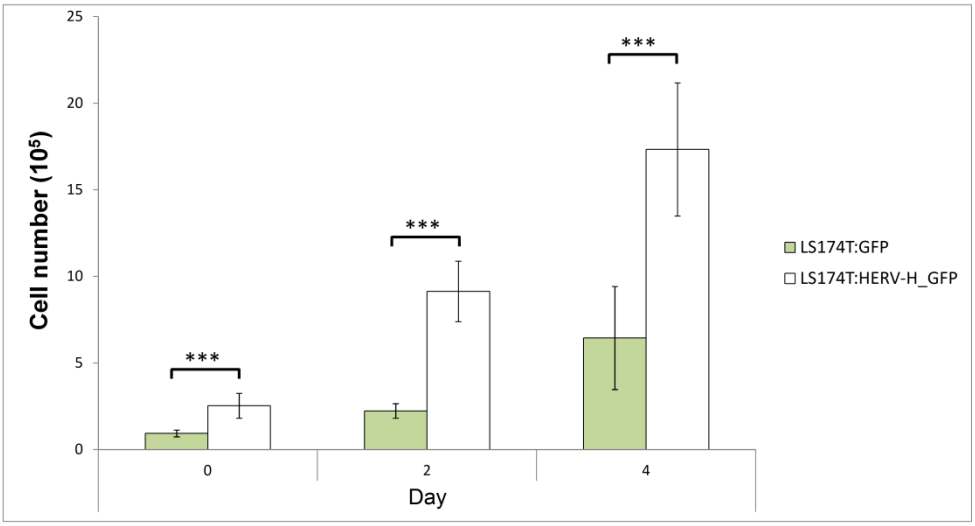
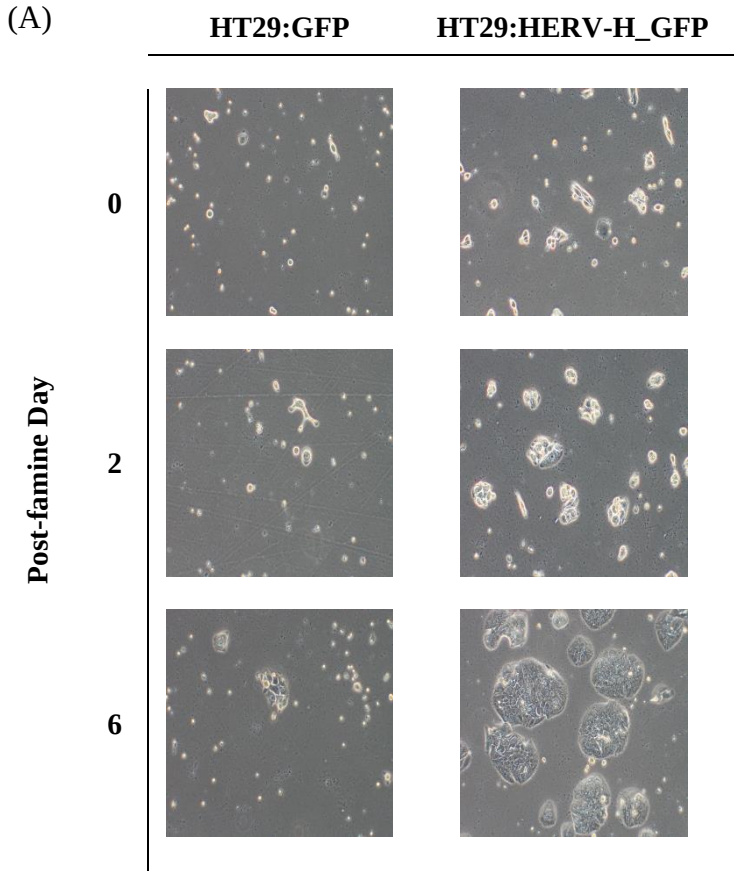
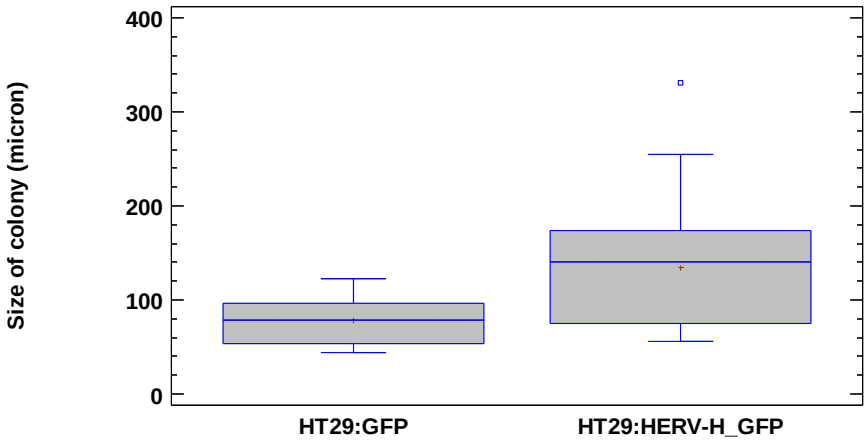


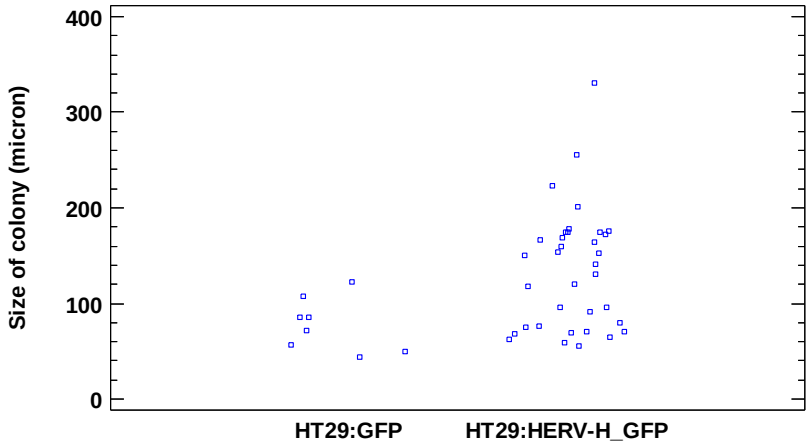
Figure 3.10. Proliferation rate of transfected LS174T cells following 30-day of serum deprivation. Experiments were done in triplicates. ***, *P* <0.001.



(B)



(C)



Comparison of Medians
Median of sample 1: 78.5
Median of sample 2: 141.0

Mann-Whitney (Wilcoxon) W-test to compare medians
Null hypothesis: median1 = median2
Alt. hypothesis: median1 NE median2

Average rank of sample 1: 12.125
Average rank of sample 2: 24.2571

W = 219.0 P-value = 0.0142807
Reject the null hypothesis for alpha = 0.05.

Figure 3.11. Survival of transfected HT29 cells following 30-day of serum deprivation. Representative images of respective transfected HT29 at indicated time points are shown. Photomicrographs (X10 magnification) were taken by IX70 microscope (Olympus) equipped with DP2-BSW software (Olympus). (B) Box-and-Whisker plot showing 25th, 50th and 75th percentiles (horizontal lines) of various sizes of colonies. The plus sign (+) indicates the location of the sample mean. (C) Dot plot showing the spread of colony sizes. Mann-Whitney W-test indicates the ability of HT29:HERV-H_GFP to form numerous large colonies is significantly different (P < 0.05) from that of HT29:GFP.

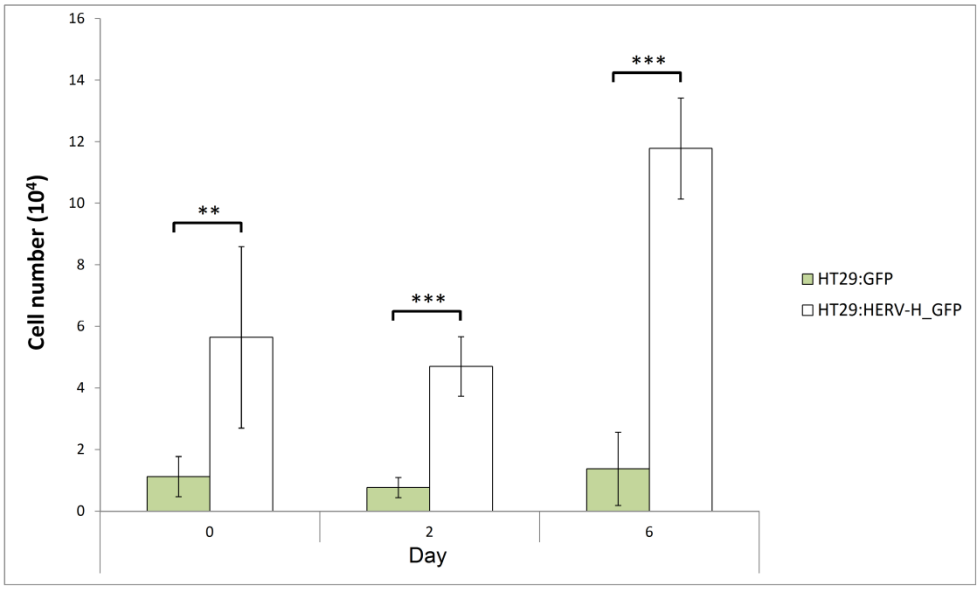


Figure 3.12. Proliferation rate of transfected HT29 cells following 30-day of serum deprivation. Experiments were done in triplicates. **, P <0.01; ***, P <0.001.

3.6.7 CD133 and CD44 expression in HERV-H transfected HT29 and LS174T cells.

CD133 (Ricci-Vitiani et al., 2007) and CD44 (Dalerba et al., 2007) are widely used as markers of cancer stem cells. To examine the effect of HERV-H on the expression of these cancer stem cell markers, cytometric analysis

was performed to analyse the expression profiles of these CD markers in the transfected HT29 and LS174T cell lines. The relative percentages of cells expressing CD133 and CD44 in transfected HT29 are presented in Figure 3.14 and 3.15 respectively. Similarly, the relative percentages of cells expressing CD133 and CD44 in transfected LS174T cells are presented in Figure 3.17 and 3.18 respectively. Generally, up-regulation of HERV-H expression augments the expression levels of CD133 and CD44 in both cell lines, and all augmentation are statistically significant (P <0.05).

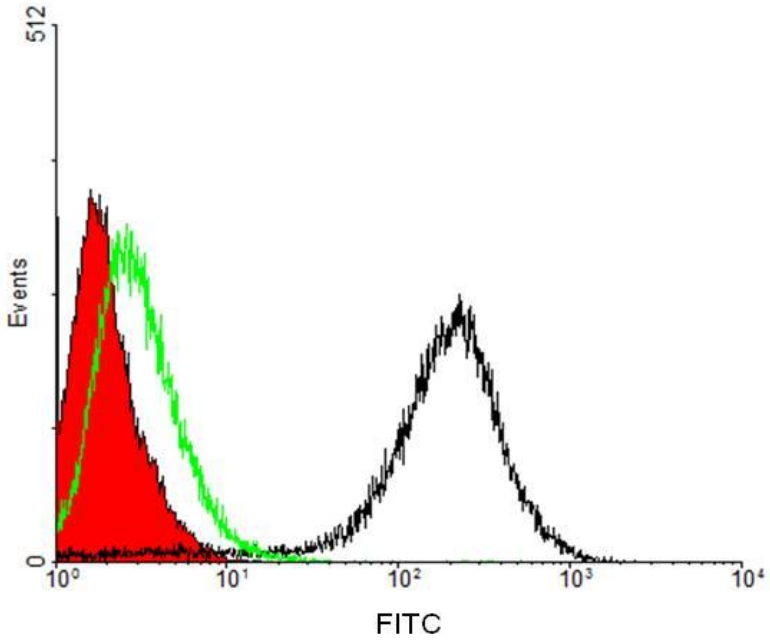
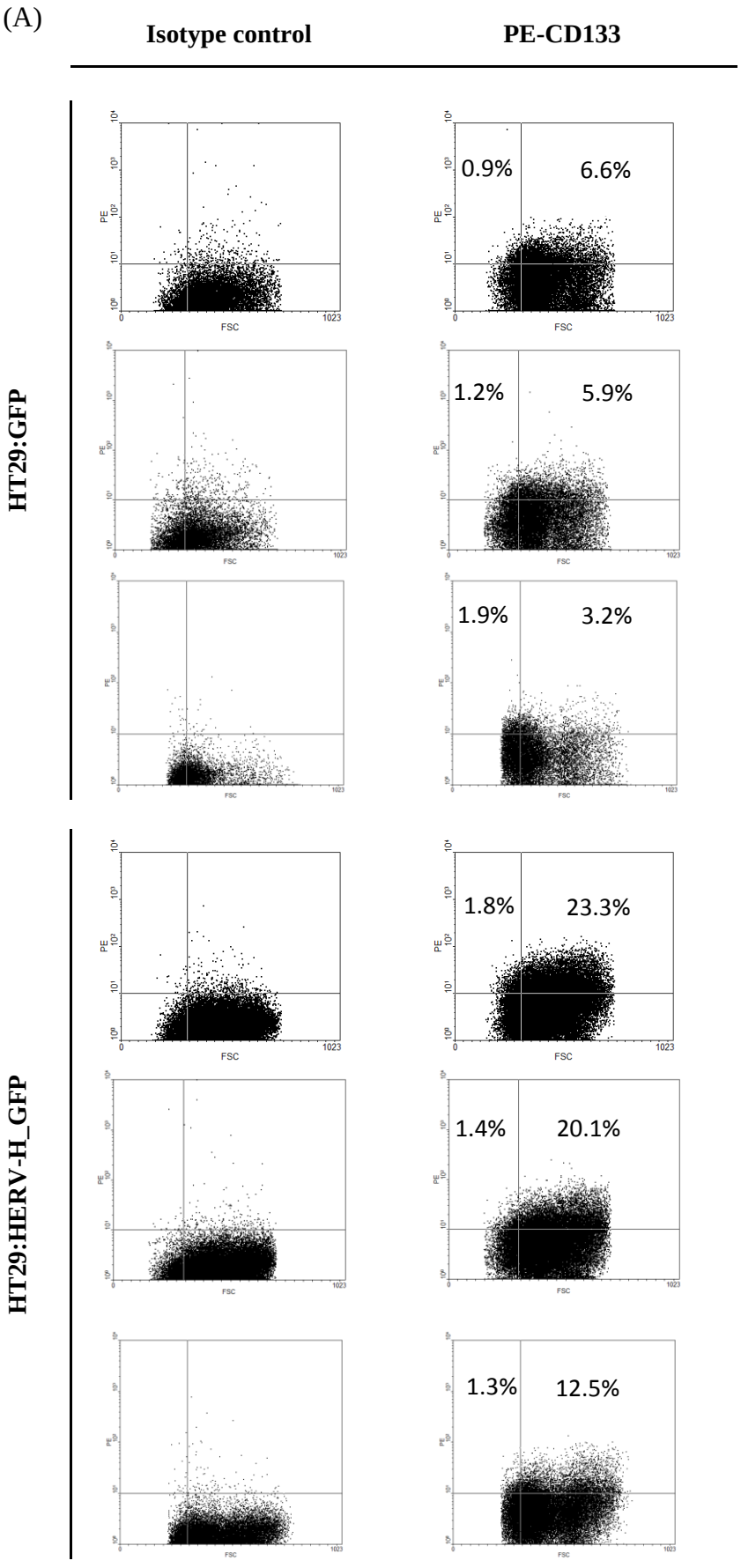


Figure 3.13. Cytometric analysis of HT29 cells. The red histogram indicates untreated HT29 cells, green and black lines indicate HT29:HERV-H_GFP and HT29:GFP cells respectively

(A)


Isotype control
PE-CD133
HT29:GFP
HT29:HERV-H_GFP
0.9%
6.6%
1.2%
5.9%
1.9%
3.2%
1.8%
23.3%
1.4%
20.1%
1.3%
12.5%
PE
FSC
0
1023

(B)

Setting	HT29:GFP		HT29:HERV-H_GFP	
	CD133 +	CD133 –	CD133 +	CD133 –
Expt A	7.5	92.5	25.1	74.9
Expt B	7.1	92.9	22.3	77.7
Expt C	5.1	94.9	13.8	86.2

(C)

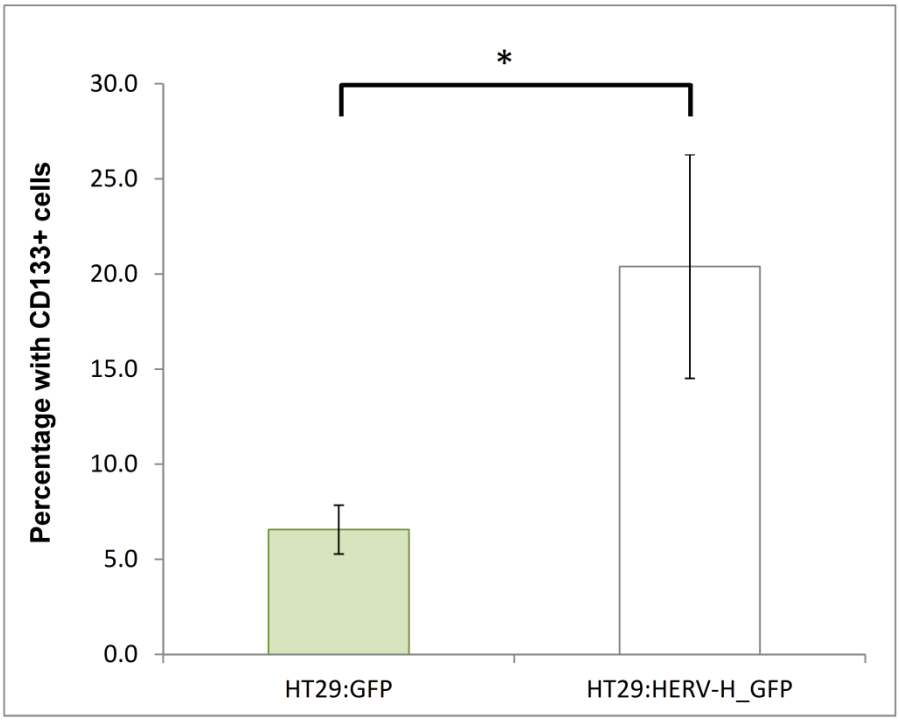
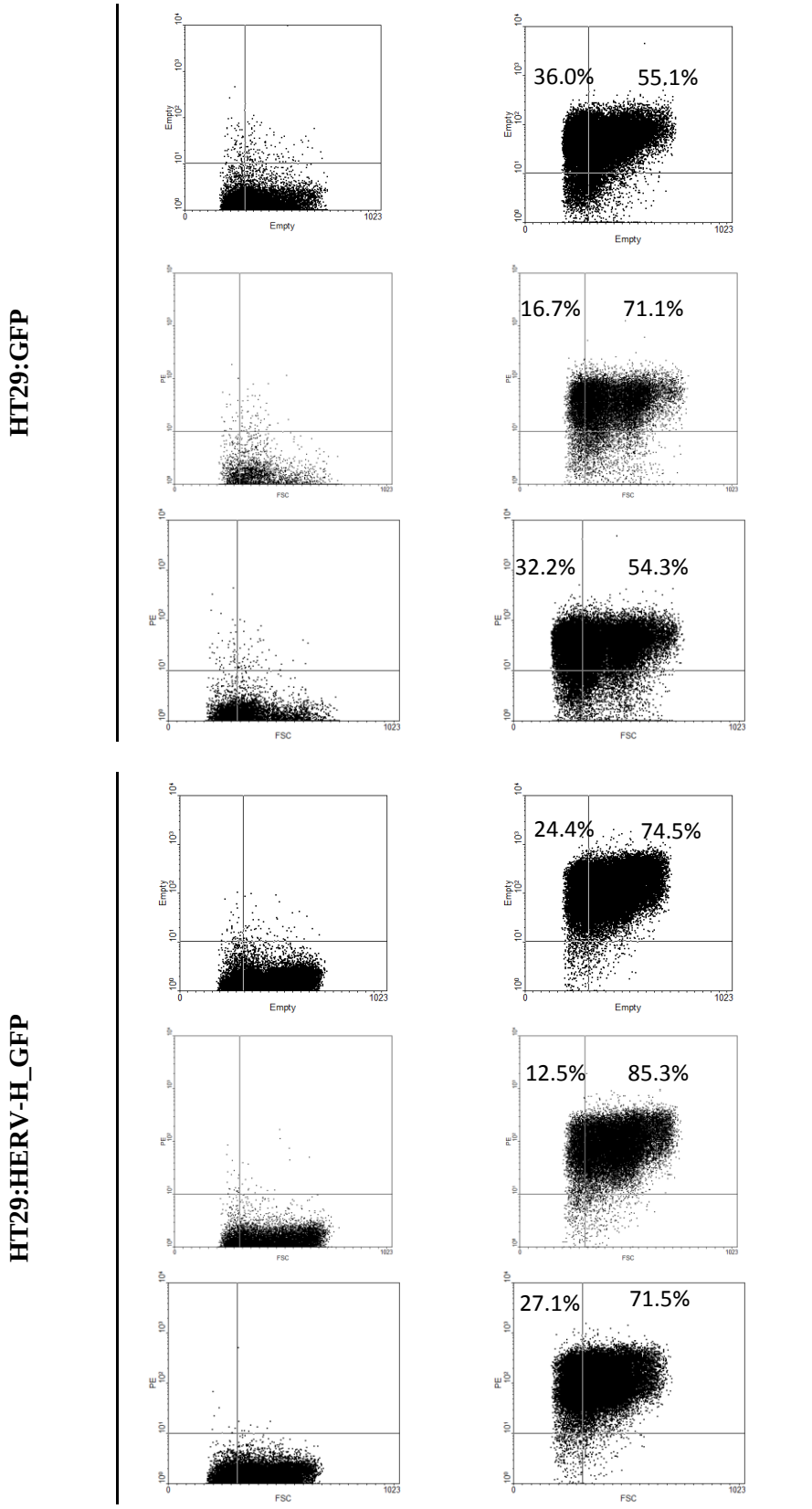


Figure 3.14. Cytometric analysis of CD133 on HT29 cells. (A) Isotypic controls were used to establish the right gating. The top portion (left upper quadrant and right upper quadrant) is CD133 positive. (B) Percentage of CD133 positive and negative cells in transfected HT29 cells. (C) Graph showing the significant difference in CD133 expression levels (P = 0.035). Independent triplicate experiments are shown. *, P <0.05.

(A)


Isotype control
PE-CD44
HT29:GFP
HT29:HERV-H_GFP
36.0%
55.1%
16.7%
71.1%
32.2%
54.3%
24.4%
74.5%
12.5%
85.3%
27.1%
71.5%
Empty
PE
FSC
0
1023

(B)

Setting	HT29:GFP		HT29:HERV-H_GFP	
	CD44 +	CD44 –	CD44 +	CD44 –
Expt H1	91.1	8.9	98.9	1.1
Expt H2	87.8	12.2	97.8	2.2
Expt H3	86.4	13.6	98.6	1.4

(C)

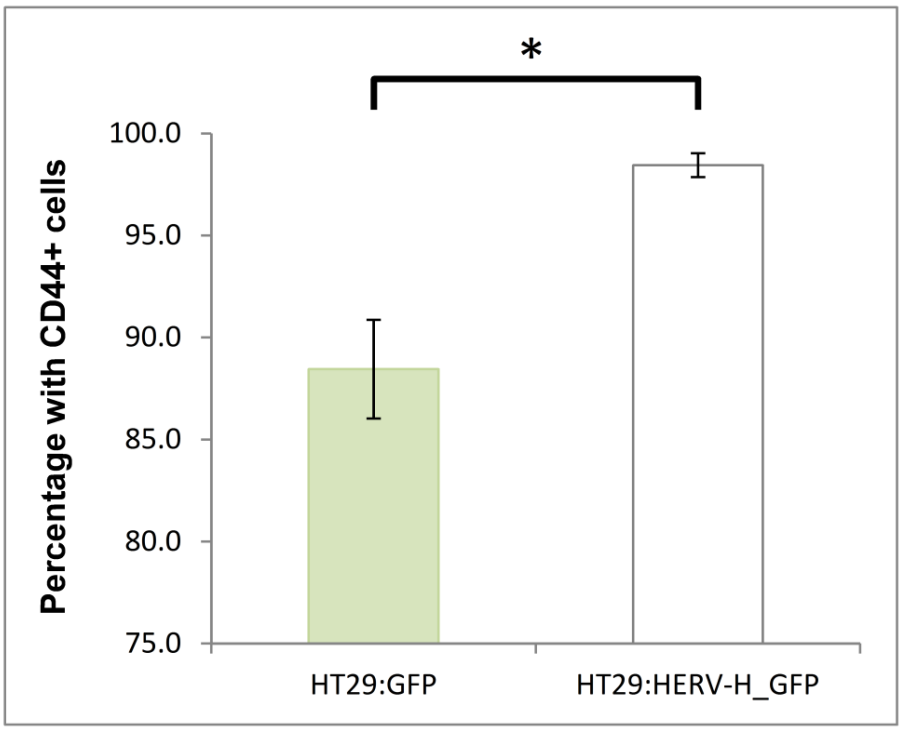


Figure 3.15. Cytometric analysis of CD44 on HT29 cells. (A) Isotypic controls were used to establish the right gating. The top portion (left upper quadrant and right upper quadrant) is CD44 positive. (B) Percentage of CD44 positive and negative cells in transfected HT29 cells. (C) Graph showing the significant difference in CD44 expression levels (P = 0.016). Independent triplicate experiments are shown. *, P <0.05.

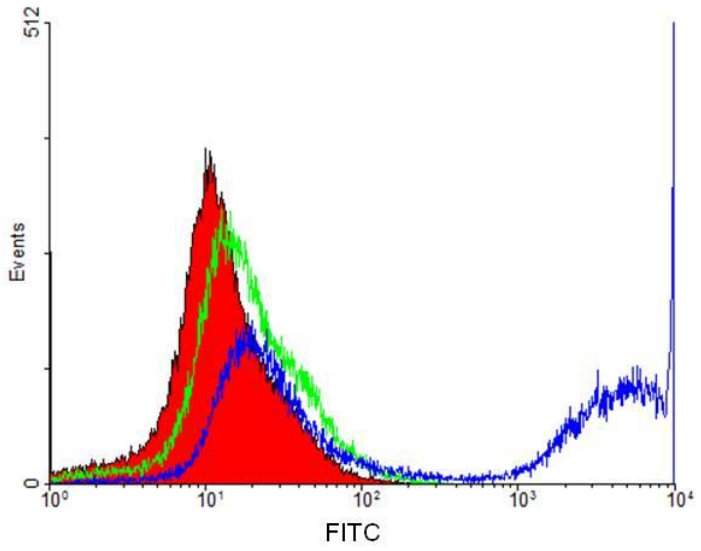
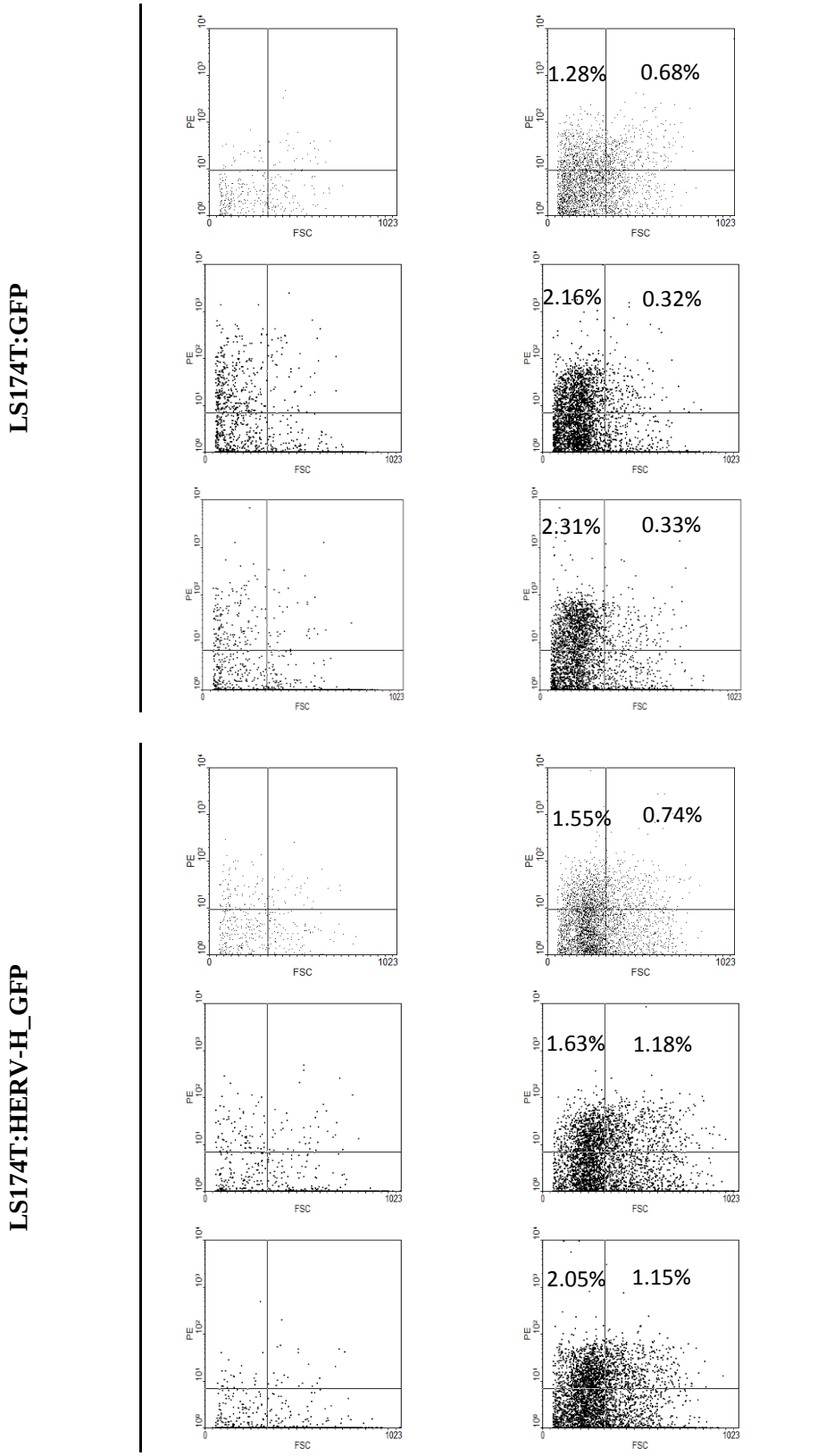


Figure 3.16. Cytometric analysis of LS174T cells. The red histogram indicates untreated LS174T cells, green and blue lines indicate LS174:HERV-H_GFP and LS174T:GFP cells respectively.

(A)


Isotype control
PE-CD133
LS174T:GFP
LS174T:HERV-H_GFP
1.28%
0.68%
2.16%
0.32%
2.31%
0.33%
1.55%
0.74%
1.63%
1.18%
2.05%
1.15%
PE
FSC
0
1023

(B)

Setting	LS174T:GFP		LS174T:HERV-H_GFP	
	CD133 +	CD133 –	CD133 +	CD133 –
Expt A	1.96	98.04	2.29	97.71
Expt B	2.48	97.52	2.81	97.19
Expt C	2.64	97.36	3.20	96.80

(C)

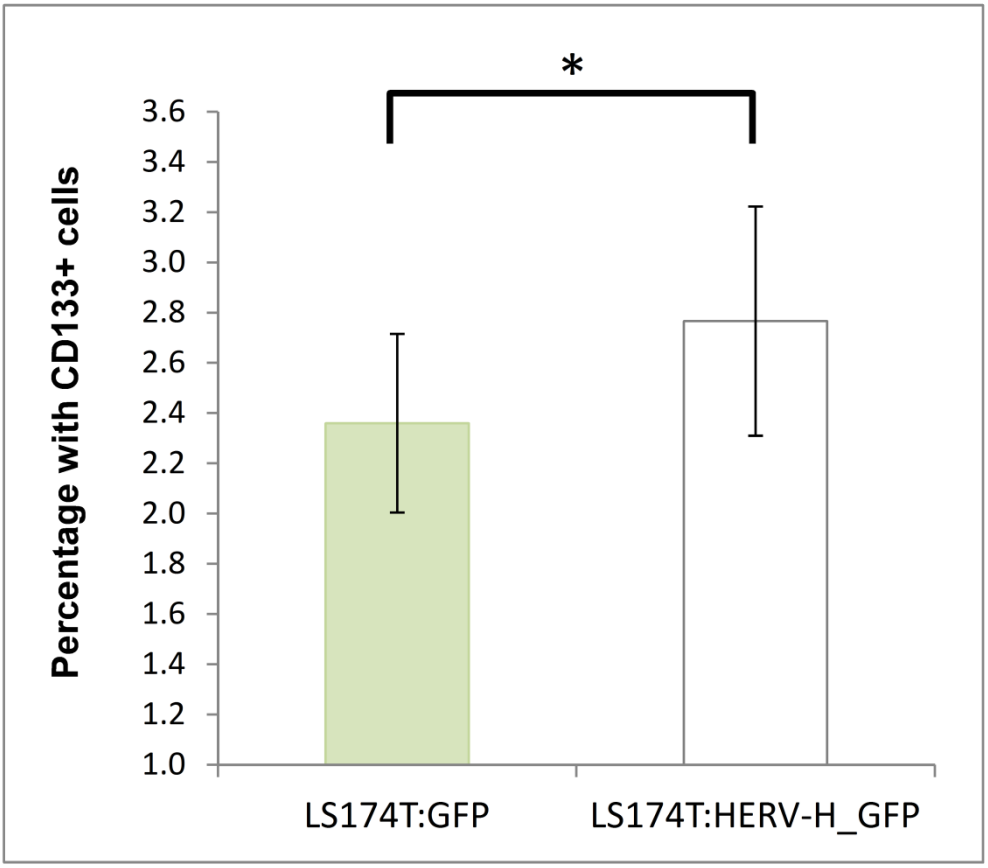
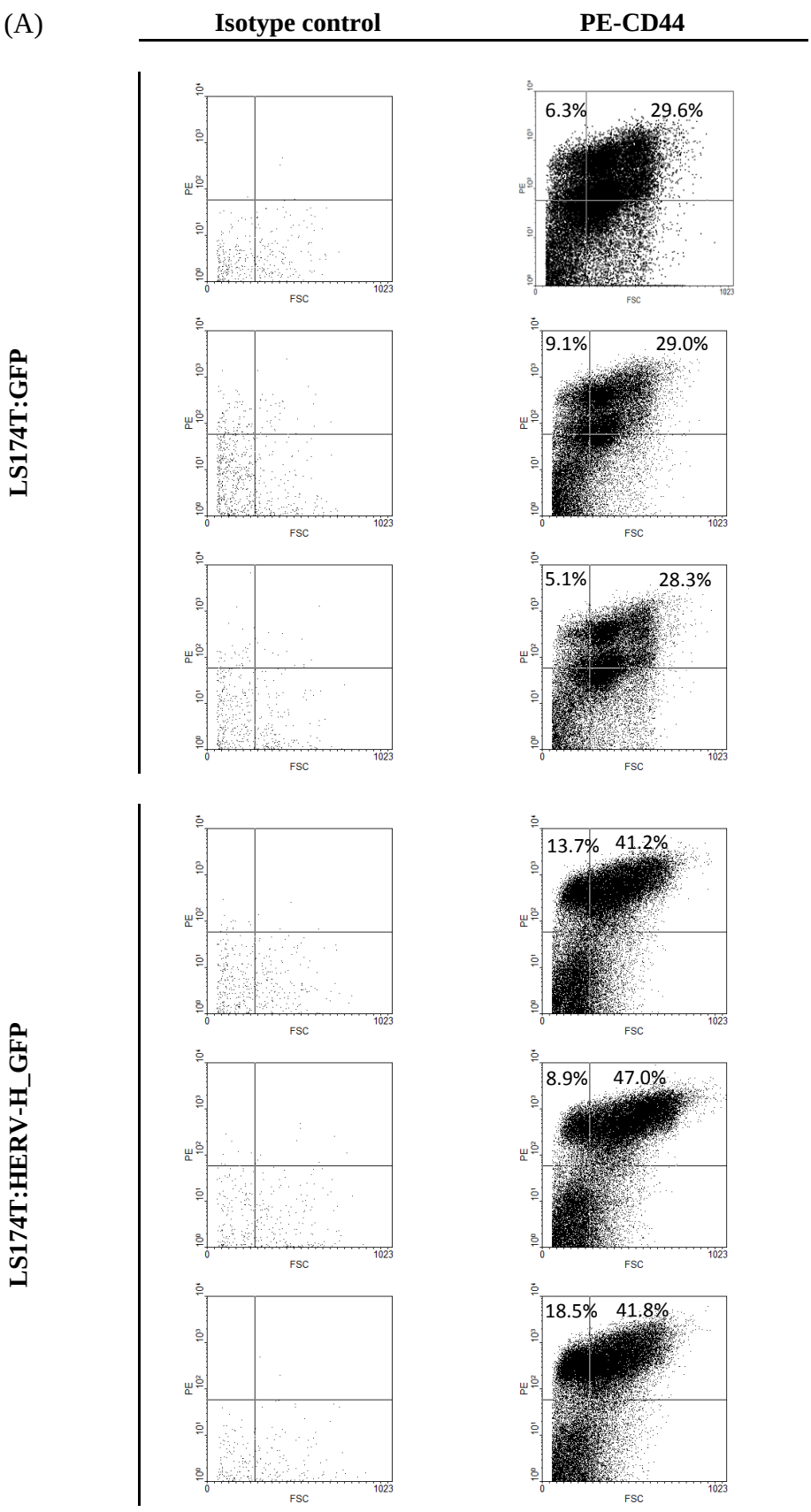


Figure 3.17. Cytometric analysis of CD133 on LS174T cells. (A) Isotypic controls were used to establish the right gating. The top portion (left upper quadrant and right upper quadrant) is CD133 positive. (B) Percentage of CD133 positive and negative cells in transfected LS174T cells. (C) Graph showing the significant difference in CD133 expression levels (P = 0.017). Independent triplicate experiments are shown. *, P <0.05.


(A)
Isotype control
PE-CD44
LS174T:GFP
LS174T:HERV-H_GFP
6.3%
29.6%
9.1%
29.0%
5.1%
28.3%
13.7%
41.2%
8.9%
47.0%
18.5%
41.8%
PE
FSC
0
1023

(B)

Setting	LS174T:GFP		LS174T:HERV-H_GFP	
	CD44 +	CD44 –	CD44 +	CD44 –
Expt A	35.9	64.1	54.9	45.1
Expt B	38.1	61.9	55.9	44.1
Expt C	33.4	66.6	60.3	39.7

(C)

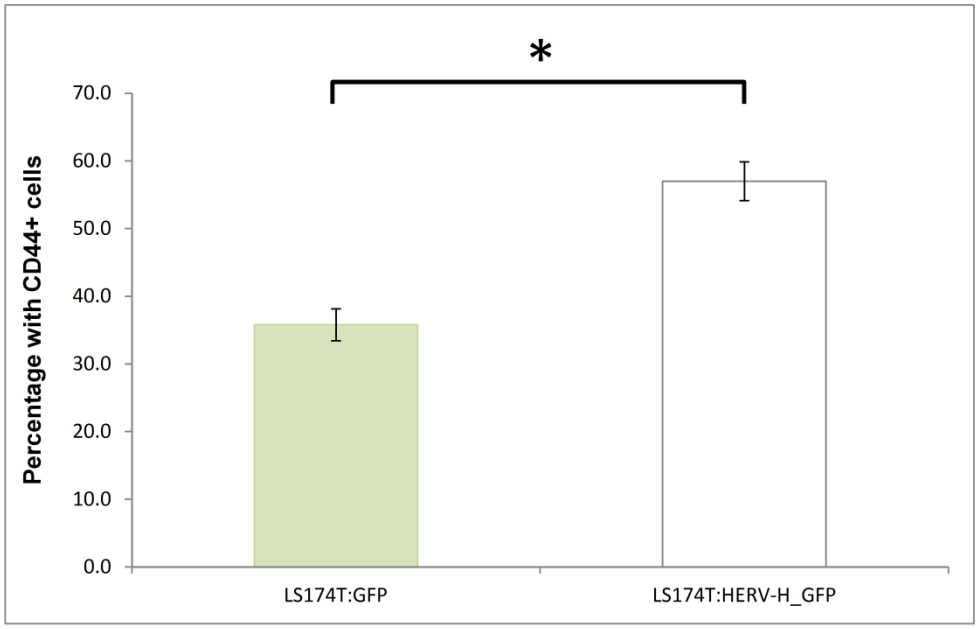
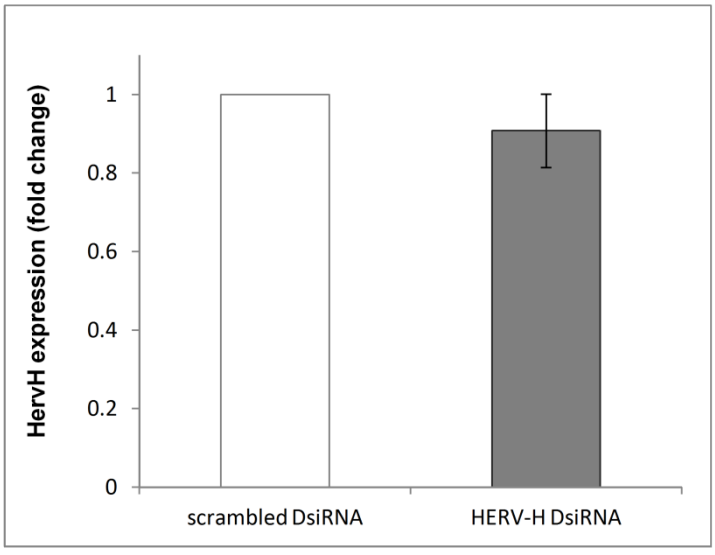


Figure 3.18. Cytometric analysis of CD44 on LS174T cells. (A) Isotypic controls were used to establish the right gating. The top portion (left upper quadrant and right upper quadrant) is CD44 positive. (B) Percentage of CD44 positive and negative cells in transfected LS174T cells. (C) Graph showing the significant difference in CD44 expression levels (P = 0.018). Independent triplicate experiments are shown. *, P <0.05.

3.6.8 Inhibition of HERV-H gene expression and cell proliferation by DsiRNA

To address the reversal effect of DsiRNA on increased proliferation rate of cells transfected with HERV-H, cells transfected with pEGFP-N2:HERV-H or pEGFP-N2 were selected for knockdown experiments. Equal numbers (2×10^5) of LS174T:HERV-H_GFP and LS174T:GFP cells were transfected with DsiRNA directed against HERV-H. Quantitative RT-PCR was performed to determine the effect of HERV-H targeting DsiRNA on the expression of HERV-H in transfected LS174T cells. The results obtained showed a significant (60%) down-regulation of HERV-H expression in cells treated with HERV-H:DsiRNA when compared to cell treated with scrambled DsiRNA (Figure 3.19). Further examination on the effect of DsiRNA on cell viability was conducted after 72 hours of DsiRNA treatment. The results in Figure 3.20 clearly indicated that the knockdown of HERV-H had resulted in a marked inhibition of cell proliferation (at 10 nM; P <0.001).

(A)



(B)

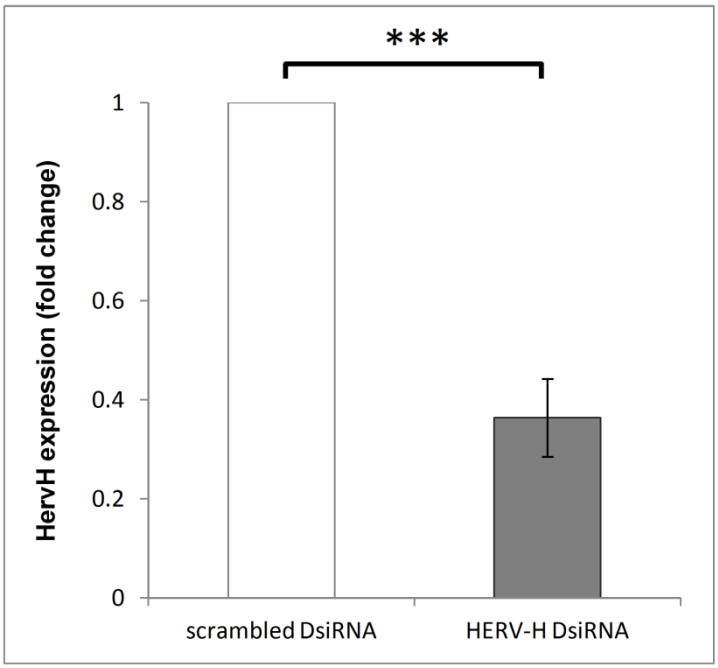
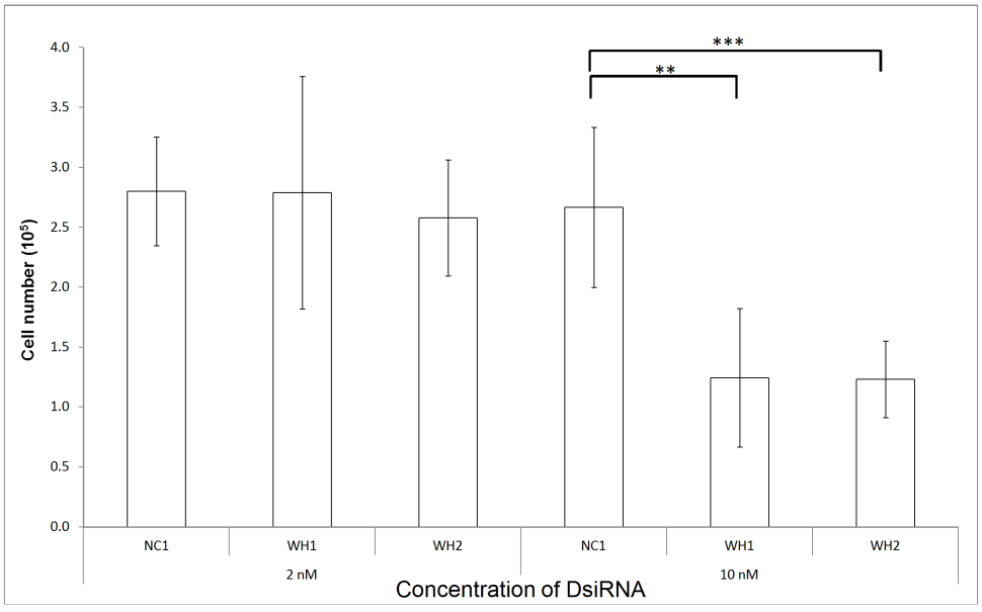


Figure 3.19. Reduction of HERV-H expression in LS174T cells by DsiRNA. After 72 hours of DsiRNA treatment, RNA was extracted and the levels of HERV-H mRNA were determined by quantitative RT-PCR. LS174T:GFP cells were used as the control. The rate of HERV-H expression was normalized to the expression of β-actin. Scrambled DsiRNA refers to the negative control group and HervH DsiRNA refers to the HERV-H silencing group. Relative gene expression was quantified based on $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001). (A) LS174T:GFP cells (B) LS174T:HERV-H_GFP cells. ***, P <0.001.

(A)



(B)

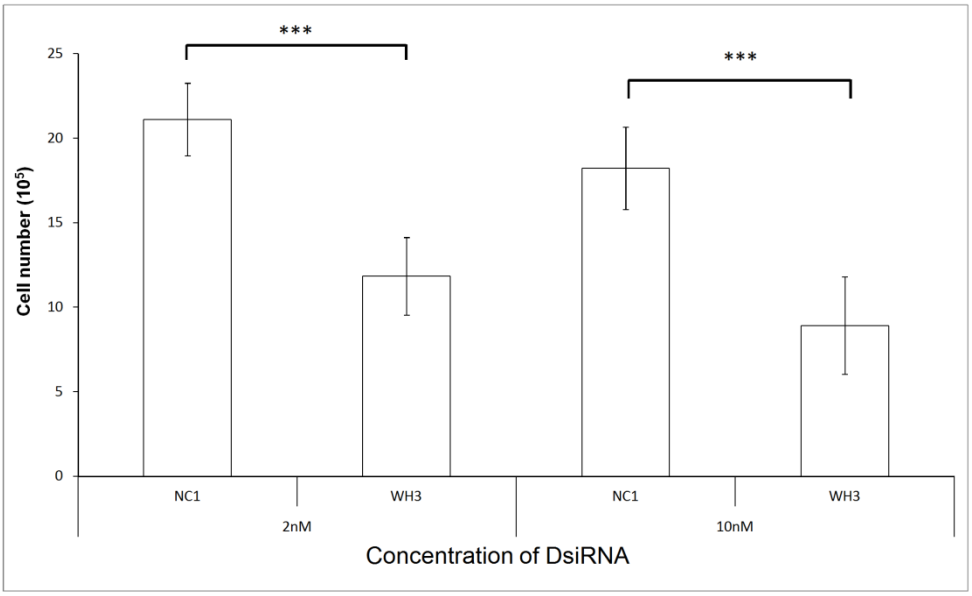
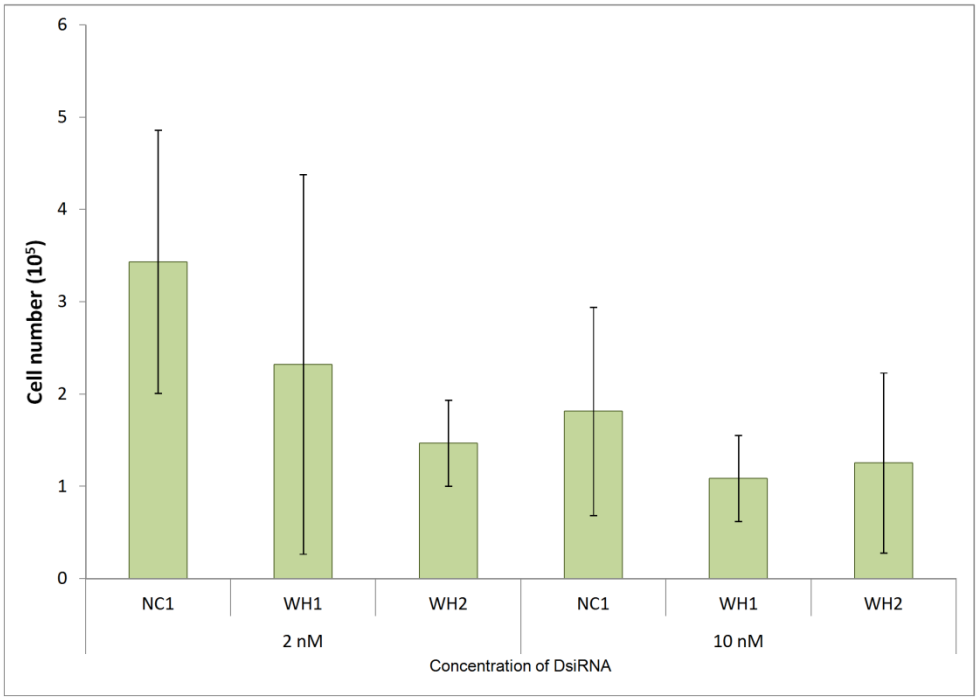


Figure 3.20. Effect of DsiRNA on cell proliferation of LS174T:HERV-H_GFP. Seventy two hours after DsiRNA transfection, the effect of various DsiRNA (WH1, WH2 and WH3) at 2nM and 10 nM on cell proliferation was determined by trypan blue assay. Cells treated with NC1, scrambled DsiRNA, served as the negative control. WH1, WH2 & WH3 refer to DsiRNAs targeting at HERV-H and these were the HERV-H silencing experiment groups. Experiments were done in triplicates. (A) Effect of NC1, WH1 and WH2 (B) Effect of NC1 and WH3. **, *P* <0.01; ***, *P* <0.001.

(A)



(B)

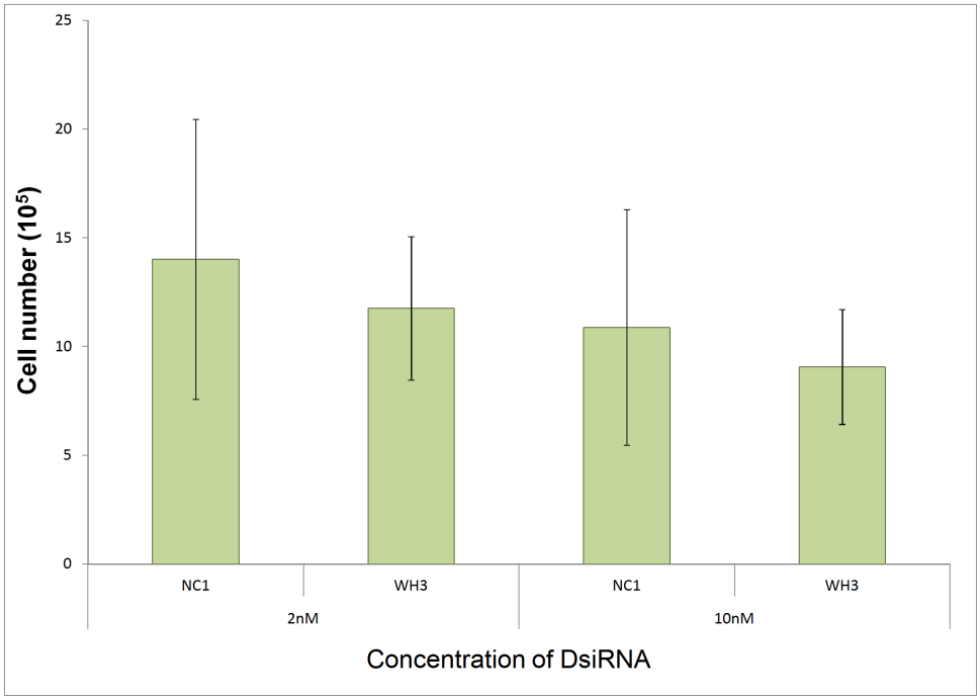


Figure 3.21. Effect of DsiRNA on cell proliferation of LS174T:GFP cells. Seventy two hours after DsiRNA transfection, the effect of various DsiRNA (WH1, WH2 and WH3) at 2nM and 10 nM on cell proliferation was determined by trypan blue assay. Cells treated with NC1, scrambled DsiRNA, served as the negative control. WH1, WH2 & WH3 refer to DsiRNAs targeting at HERV-H and these were the HERV-H silencing experiment groups. Experiments were done in triplicates. (A) Effect of NC1, WH1 and WH2 (B) Effect of NC1 and WH3.

3.7 Discussion

To date, little is known about the role of endogenous retroviruses in cancer development. Although many studies have shown correlation of the transcriptional activities of endogenous retroviral elements with tumour specimens, the precise molecular pathogenesis is still poorly understood.

Nevertheless, endogenous retroviruses have generated much interest in the clinical research fields. The involvement of endogenous retroviral

elements in the study of multiple sclerosis is progressing with interesting clinical data (Nissen et al., 2013). These include the strong association of HERV-K18.3 haplotype (de la Hera et al., 2013, Tai et al., 2008) or HERV-Fc1 (Nexo et al., 2011) with multiple sclerosis susceptibility, and a gender difference in HERV-W copy numbers in multiple sclerosis patients (Garcia-Montojo et al., 2013). Review articles summarizing the advances and controversial issues on neuropathogenesis (Hon et al., 2013, Christensen, 2010, Antony et al., 2011) are also available. Conversely, endogenous retroviruses have also been implicated in autism spectrum disorders (Balestrieri et al., 2012) as well as severe psychiatric disorders like schizophrenia and bipolar disorder (Perron et al., 2012b, Perron et al., 2008, Frank et al., 2005, Karlsson et al., 2001).

Similarly, there is an increasing interest in the study of HERVs as causal factors in cancer biology (Lower et al., 1993, Herbst et al., 1996, Wang-Johanning et al., 2003). Activation of HERV-K is found to be necessary for the malignant progression of melanoma (Serafino et al., 2009). The expression of HERV-K gag antigen and the presence of anti-HERV-K-gag antibodies have been demonstrated to be correlated to prostate cancer progression (Reis et al., 2013). Still, the quest for the aetiopathogenic role of endogenous retroviral elements in human tumours is ongoing. A simple reason for this is that their mechanistic role in disease development is still not clearly defined.

The six hallmarks of cancer have been proposed by Douglas Hanahan and Robert A. Weinberg in 2000. These six acquired capabilities are depicted in Figure 3.22 (Colotta et al., 2009). This widely regarded conceptual framework forms the basis of cancer biology and sets the tone of findings discussion in this study.

3.7.1 Proliferation: Heightened replicative potential observed

In a series of experiments conducted, our results have demonstrated the oncogenic potential of a 93-amino acid HERV-H peptide. HERV-H overexpressing cells were proliferating at a higher rate. This corroborates one of the six hallmarks of cancer – limitless replicative potential. Although arguably the replicative activity of colorectal cancer cells should have already been developed in the established cell lines, it is the increased rate of proliferation that set the underlying mechanism for expansive tumour growth (Cho et al., 2006, Csibi et al., 2013).

In the experiments in which HERV-H was knocked down by dicer-substrate small interfering RNA the cells were observed to decrease their growth rate with concomitant decrease in HERV-H expression. These data therefore illustrate a central role of HERV-H in the replicative potential of cells and clearly demonstrate that HERV-H expression is associated with the phenotypic transformation of colorectal cancer.

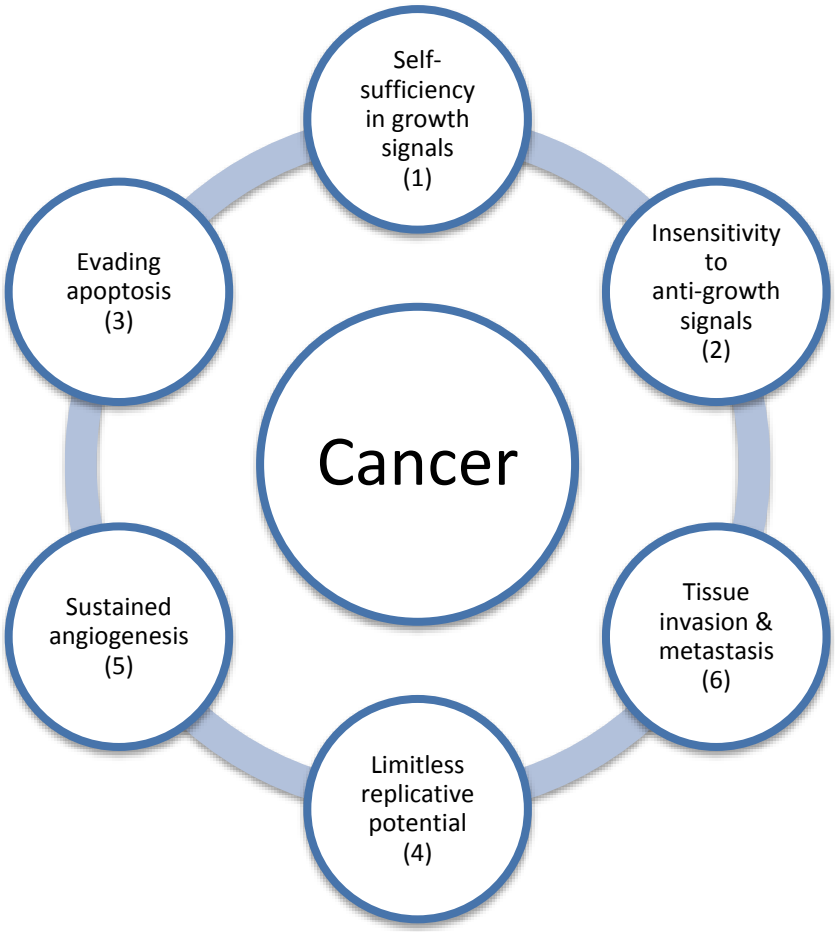


Figure 3.22. The six hallmarks of cancer proposed by Hanahan and Weinberg in 2000. The hallmarks of cancer consist of six biological capabilities acquired during the transformation of normal cells to cancer cells. The hallmarks underlie an organizing principle for rationalizing the complexities of neoplastic disease. The hallmarks are (1) sustaining proliferative signalling, (2) evading growth suppressors, (3) resisting cell death, (4) enabling replicative immortality, (5) inducing angiogenesis, and (6) activating invasion and metastasis.

3.7.2 Migration: Initial step for tissue invasion & metastasis engaged

Cell motility or the migration process of adherent cells from one location to another is the initial phase of the complex metastasis process. Cell migration can take place either as single cells or in small groups, thereby aiding the spread of cancer cells from the initial site of tumour growth to other

locations. In order for the cell to migrate, it must acquire an invasive phenotype that is characterized by both the loss of cell-cell interactions and increased cell motility (Palmer et al., 2011). In our study, HERV-H overexpressing cells were able to migrate at a higher rate. The engagement of the initial process of invasion and metastasis is therefore demonstrated. However the invasive property is not demonstrated. While the discordant effects of HERV-H on the invasion and metastasis is not clearly understood, it may be associated with the rise in CD133/CD44 expression, which will be explained later.

3.7.3 Sphere-forming ability: stem cell potential evaluated

In 2011, Douglas Hanahan and Robert A. Weinberg put forward four more hallmarks of cancer, namely, (1) deregulation of cellular energetic, (2) avoidance of immune destruction, (3) genome instability and mutation, and (4) tumour promoting inflammation (Figure 3.23). Interestingly, it is in this new proposal that the role of cancer stem cells found in the tumour microenvironment was discussed (Hanahan and Weinberg, 2011). Undeniably, it is the presence of cancer stem cells that may account for the sudden termination of tumour dormancy. On the other hand, it is regarded that the sphere-forming assay evaluate the potential of a cell to behave as a stem cell (Pastrana et al., 2011). With these notions in mind, the large sphere-forming ability gained by the HERV-H overexpressing LS174T cells provided the capability of initiating and sustaining tumour growth by forming a population of cancer stem cells and descendant cells. Obviously, this presents a poor

clinical prognosis and offers limited effective therapeutic modalities (Kleffel and Schatton, 2013).

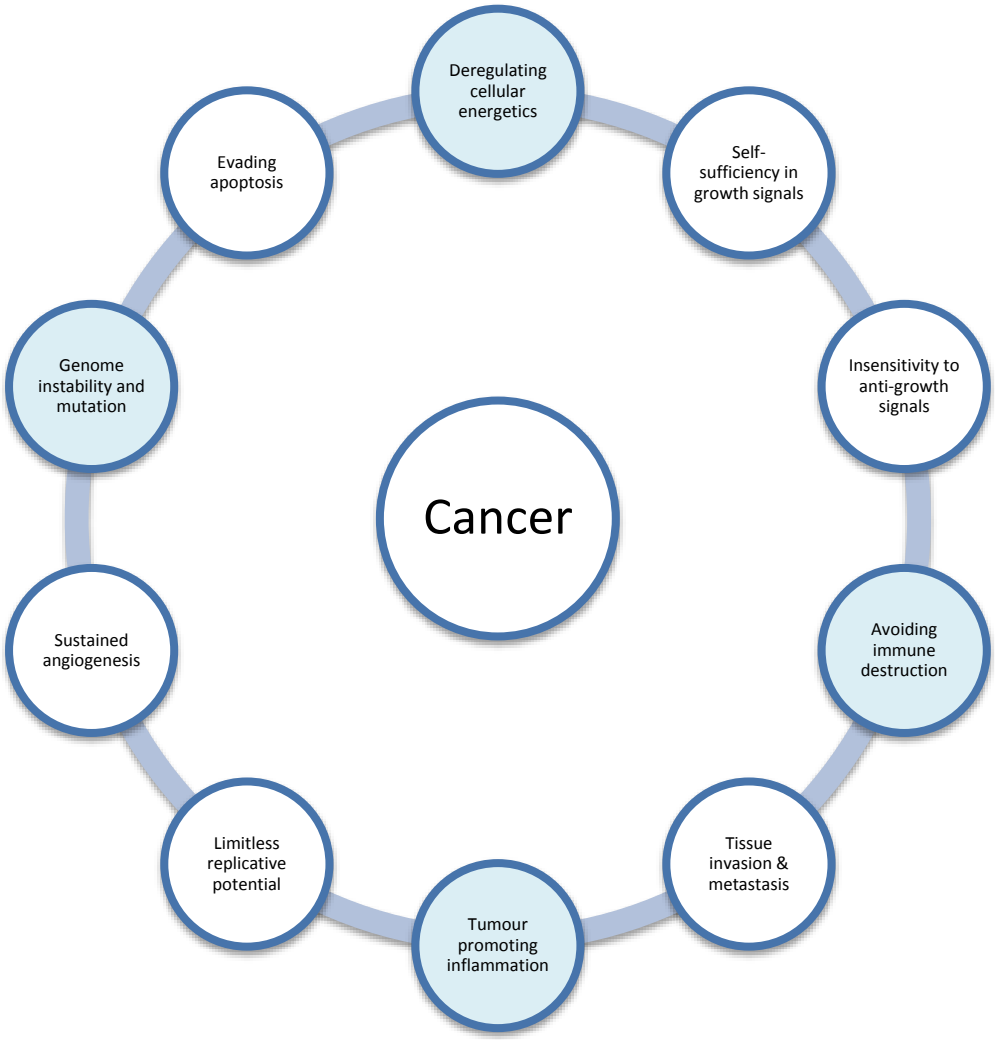


Figure 3.23. The next-generation hallmarks of cancer proposed by Hanahan and Weinberg in 2011. The progress in cancer research over the past decade has provided many interesting developments in the field. Consequently, four new hallmarks (shaded circles), viz. deregulation of cellular energetic, avoidance of immune destruction, genome instability and mutation, and tumour promoting inflammation were added. Along with these new hallmarks, the concept of "tumour environment" was introduced, and the signalling networks in the tumour microenvironment during malignant progression were discussed.

The sphere-forming ability of colon cancer cells have also been correlated with the expression of CD133 (Ricci-Vitiani et al., 2007) and CD44 (Su et al., 2011). Several studies stipulated the experimental conditions for generating cancer-stem-cells-like sphere cells from colorectal cancer cells (Todaro et al., 2007, Kanwar et al., 2010, Fan et al., 2011, Hwang et al., 2011), which include the use of stem cell medium or serum-free medium supplemented with various mix of B27, epidermal growth factor, fibroblast growth factor and insulin. Our study revealed the swift gain of sphere-forming ability in HERV-H transfected LS174T cells, without the need of special experimental conditions as previously stated. This suggests that the HERV-H element might act, to some extent, as a molecular switch/link to a myriad of stem cell potency factors, resulting in the quick formation of cancer-stem-cell-like sphere cells.

3.7.4 Stem cell phenotype: CD133 and CD44 highly expressed

CD133, a five-transmembrane glycoprotein of 120 kDa molecular weight, was first demonstrated in haematopoietic stem and progenitor cells (Yin et al., 1997, Miraglia et al., 1997). When tumour initiating characteristic was found to be common in CD133-positive cells isolated from many neoplasms (Yin et al., 1997, Singh et al., 2004, Collins et al., 2005, Suetsugu et al., 2006, Hermann et al., 2007, Ferrandina et al., 2008, O'Brien et al., 2007), it was subsequently widely used to identify and isolate stem cells and cancer stem cells. CD44, a transmembrane glycoprotein which functions in cell adhesion, migration, homing, proliferation, survival and apoptosis, is regarded as a cancer stem cell marker for many neoplasms (Zoller, 2011, Du

et al., 2008). As CD133 and CD44 are widely used markers of cancer stem cells, their clinical relevance in prognostication is slowly gaining its traction through clinical studies (Horst et al., 2009b, Du et al., 2008, Galizia et al., 2012).

Our study has demonstrated that overexpression of HERV-H has augmented the expression of CD133 and CD44 in both HT29 and LS174T cells significantly ($P < 0.05$). In addition, HERV-H overexpressing colorectal cells were able to proliferate faster. This is consistent with the previous study (Ieta et al., 2008) in which the tumour-forming ability and higher proliferative capacity of CD133+ cells in colon cancer cell lines was demonstrated.

As previously mentioned, HERV-H overexpressing colorectal cells were unable to display the characteristics of cell invasiveness.The increased levels of CD133 in both transfected HT29 and LS174T cells did not associate with a higher degree of cell invasiveness. Previously, CD133+ HT29 cells have been shown to possess higher invasive abilities when compared with CD133– HT29 cells (Ieta et al., 2008). Similarly, CD133 was demonstrated to play an important role in HCT116 colon cancer cell invasion (Zhang et al., 2013a). On the other hand, controversial findings have reported that CD133 is unlikely to contribute functionally to the metastatic phenotype of colon cancer cells (Horst et al., 2009c, Shmelkov et al., 2008). With these discordant findings, it may appear that CD133 may possess a dualistic nature.

Similarly, greater cell invasiveness is not observed in CD44-positive HERV-H overexpressed HT29 and LS174T cells. While CD44-overexpressing SW480 cells colon cancer cells have been shown to exhibit a 2.5-fold increase in cell invasion (Cho et al., 2012), CD44 overexpression has also been demonstrated to inhibit invasion *in vitro* for prostate cancer (Yang et al., 2010). While this may arguably be taken as the metastatic role of CD44 is tissue-specific, previous studies have reported that a reduced expression of CD44 was found in colon cancer metastases (Weg-Remers et al., 1998, Harada et al., 2001). Of interest, a recent study has shown that the expression of CD44 by LS174T colon carcinoma cells functions as a tumour suppressor (Dallas et al., 2012) and that the knockdown of CD44 exhibits a large increase in metastatic potential. Taken together, CD44 may be regarded as the mediator for decreased cell invasiveness.

3.7.5 Serum independence: Self-sufficiency in growth signals met and possible reprogrammed cellular energetics

Normal cells need mitogenic growth signals to proliferate. In other words, growth promoting signals are required to direct the cells through cycles of cell growth and division. On the other hand, cancer cells are able to acquire the mitogenic signalling they need to sustain proliferation. This may take place in an autocrine manner, leading to overexpression of growth receptors or constitutive activation of signalling pathways (Hanahan and Weinberg, 2000, Hanahan and Weinberg, 2011).

One of the most interesting findings in this study is the increased revival rate of HERV-H-overexpressing LS174T and HT29 cells after 30 days of serum deprivation. Fetal bovine serum contains high levels of growth stimulatory factors. Hence, supplementation of fetal bovine serum in the culture medium always helps in the optimisation of *in vitro* cell culture (Zheng et al., 2006). With prolonged serum deprivation, growth of normal cells is very unlikely. Although an insufficient supply in serum nutrients may not have an overt effect on cancer cell viability, starved cancer cells are sensitive to apoptosis (Braun et al., 2011). Nevertheless, the alteration of cellular metabolism to suit changes in nutrient availability suggests cancer cells have successfully adapted their metabolism to ensure survival. It is reported that such metabolic alteration in cancer cells is modulated at the epigenetic level (Yun et al., 2012). Interestingly, the Warburg effect (Warburg, 1956) has been found to play a role in this phenomenon via "aerobic glycolysis" and glutamine can be used to support the synthesis of cellular building blocks like amino acids, ribonucleotides and lipids (Dang, 2009). With growth media containing GlutaMAX™, a more stable source of glutamine, the Warburg effect may be the factor for cancer cells survival.

Here we show for the first time that HERV-H expressing LS174T and HT29 cells possess the advantage of surviving through long period of serum deprivation. The ability to pull through such a long period of serum deprivation may be associated with the increased expression of CD133. Population of CD133+ cells was shown to be capable of growing as

undifferentiated colon-spheres in a serum-free media supplemented with EGF and FGF-2 (Ricci-Vitiani et al., 2007). While the sphere-forming ability, higher proliferative capacity and survival of long period of serum-free cultivation may be attributed to the increase in CD133 expression, it should be realised that it is after all, the up-regulation of HERV-H expression that initiated the whole process.

3.7.6 Concluding remarks

Taken together, overexpression of HERV-H promoted colorectal cancer cell proliferation, motility, and sphere formation and reduced serum dependence via increased expression of CD133 and CD44. Specifically, the expression of HERV-H is conferred by the 93-aa Gag protein, which is one of the ORFs in the HERV-H sequence. As such, this study represents part of the determination of the roles of HERV-H in colorectal carcinogenesis. It is unknown if HERV-H up-regulation in colorectal cancer biopsies is a bystander effect, but results presented in this chapter appear to suggest otherwise, given that HERV-H directly regulates various hallmarks of cancer. In the next chapter, the involvement of various signalling pathways and key components underlying HERV-H's apparent oncogenic prowess will be discussed.

Role of HERV-H in viral mediated signalling

Chapter 4

CHAPTER 4

ROLE OF HERV-H IN VIRAL MEDIATED SIGNALLING

4.1 Introduction

In 2000, the six hallmarks of cancer described by Douglas Hanahan and Robert Weinberg provided a new paradigm of cancer development framework (Hanahan and Weinberg, 2000). Later in 2011, both Hanahan D and Weinberg R put forward another 4 new hallmarks to account for the advances in cancer research in the field (Hanahan and Weinberg, 2011). The notion of a multistep process of human tumour pathogenesis, which entails the accumulation of multiple independent mutations that lead to the deregulation of cell signalling pathways coupled with uncontrolled cell growth, has always been the central dogma of cancer biology.

4.2 Viral-mediated cancers

About 20% of the human cancers are linked to infectious agents, of which 15% are associated with viruses (zur Hausen, 1991, Parkin, 2006, Zur Hausen, 2009). While viruses are notorious for devastating diseases, for example, cancers, it should be appreciated poignantly that viral carcinogenesis has contributed much to the molecular study of cancer development and the fundamentals of cell signalling pathways. Nevertheless, tumour-initiating viruses were initially known to cause malignant diseases in animals. These include Rous sarcoma virus (RSV) that causes sarcoma in chickens, Harvey sarcoma virus that causes sarcoma in rats, myeloproliferative leukaemia virus

that causes acute leukaemia in mice, AKT8 virus that causes leukaemia in mice and avian leukosis virus that causes leukaemia in chickens (Rosenberg, 1997).

Although most oncogenes, which encode proteins that control cell proliferation, differentiation and apoptosis, were first recognized in retroviruses, it must be bear in mind that oncogenes from DNA viruses also play a crucial role in the development of cancer in which the cell cycle process is modulated via the retinoblastoma protein (pRb), or constitutive expression of signalling receptors (Vogt, 2012, Croce, 2008).

On the other hand, it must also be realised that not all retroviruses carry an oncogene in their genome. Retroviruses that lack an oncogene and are still able to induce tumours are collectively known as *cis*-acting retroviruses. The mechanism by which these retroviruses causes malignancy is mediated through a process known as insertional mutagenesis (Hayward et al., 1981, Mikkers and Berns, 2003, Maksakova et al., 2006). It is in this process where provirus integrates in the vicinity of a cellular oncogene and is able to function as a transcriptional regulator. This enhances or disrupts the transcriptional activities of the cellular gene, bringing upon the oncogenic potential of the cellular gene (Uren et al., 2005, Cavazza et al., 2013, Fan and Johnson, 2011). To date, there are approximately seventy proto-oncogenes activated by proviral insertion of a non-transforming retrovirus (Rosenberg, 1997), but these have been limited to non-human carcinogenesis (McLaughlin-Drubin and Munger, 2008).

4.2.1 Retroviral oncogenes

Oncogenes were first discovered in cancer-causing retroviruses (Kalland et al., 2009). To date, there are approximately 30 retroviral oncogenes of which most are associated with avian and rodent viruses (Table 4.1). The first oncogene, *v-src*, was discovered in 1911 by Peyton Rous, who identified a transforming agent, now known as Rous sarcoma virus (RSV), in a cell-free filtrate that was capable of inducing sarcomas in injected chickens (Rous, 1911, Rous, 1983). This has led to many studies of oncogenic retroviruses (Martin, 2004) with chickens and the discovery of *c-myc* from avian myelocytomatosis virus MC29 (Duesberg et al., 1977) and *erbB* from avian erythroblastosis virus (Bister and Duesberg, 1979, Lai et al., 1979). Nevertheless the studies of oncogenic retroviruses had been extended to mammals like mice, and this had subsequently led to the discovery of *ras* from Harvey and Kirsten sarcoma viruses (Shih and Weeks, 1984, Shih et al., 1979).

Table 4.1. Oncogenic viruses and associated oncogenes

Product	Oncogene	Virus
platelet-derived growth factor-β	Sis	Simian sarcoma virus
EGFR	ErbB	Avian erythroblastosis virus
thyroid hormone receptor-α.	ErbA	Avian erythroblastosis virus
GTPase	Ha-ras	Harvey sarcoma virus
GTPase	Ki-ras	Kirsten sarcoma virus
a modular signalling link	Crk	CT10 avian sarcoma virus

a signalling protein kinase	Src	Rous sarcoma virus
a signalling protein kinase	Abl	Abelson murine leukemia
a signalling protein kinase	Akt	Akt8 murine thymoma virus
a signalling protein kinase	Mos	Moloney murine sarcoma virus
a component of the AP1 complex	Jun	Avian sarcoma virus 17
a component of the AP1 complex	Fos	Finkel–Biskis–Jinkins murine sarcoma virus
a transcription factor	Myc	Avian myelocytomatosis virus MC29
lipid kinase	Pi3k	Avian sarcoma virus 16

4.3 Viral-mediated signalling pathways

One of the hallmarks of cancer – self-sufficiency in growth signals proposed by Hanahan and Weinberg, indicates that cancer cells are able to acquire the mitogenic signalling they need to sustain proliferation. Whereas this may take place via an autocrine manner, overexpression of growth receptors or simply constitutive activation of signalling pathways (Hanahan and Weinberg, 2000, Hanahan and Weinberg, 2011), growth signalling pathways may be deregulated or even reprogrammed to regulate the deranged cancer cells. The intracellular circuitry of signalling networks is both complex and intricate. Seven signalling pathways have been implicated in cancer and embryonic stem cells. These are: (1) the JAK/STAT pathway, (2) the NOTCH signalling pathway, (3) the MAP-Kinase/ERK pathway, (4) the PI3K/AKT pathway, (5) the NFkB pathway, (6) the Wnt pathway and (7) the TGFβ pathways (Dreesen and Brivanlou, 2007). Interestingly, viral-mediated signalling pathways involve several more pathways. These include: the p53

signalling pathway, the B cell receptor signalling pathway, the JAK/STAT signalling pathway and the chemokine signalling pathway (Saha et al., 2010).

In the preceding chapter, our study has demonstrated the oncogenic potential of HERV-H in a series of experiments. However, little is known about the signalling pathways that HERV-H is involved in during the transforming process. Thus, in this study, we focused on examining the pathways that HERV-H is involved in by using PCR array technology and Western blotting technique.

4.4 Materials and Methods

4.4.1 Cell lines and culture conditions

The human colorectal cancer cell lines LS174T were transfected with pEGFP-N2 or pEGFP-N2:HERV-H as previously described in Chapter 3 (Section 3.5.1). Transfected colorectal cells were routinely cultured in Minimum Essential Medium (MEM) α supplemented with 10% fetal bovine serum and 800 μg/ml of G418. Both cell lines were maintained in a humidified atmosphere of 5% CO_2 in air at 37˚C and used when in the log phase of growth.

4.4.2 PCR array

4.4.2.1 PCR array panels

The TaqMan® OpenArray® Human Signal Transduction Panel, QuantStudio™ 12K Flex (P/N 4475392) and TaqMan® OpenArray® Human

Inflammation Panel, QuantStudio™ 12K Flex (P/N 4475389) were obtained from Life Technologies Corporation, Singapore.

The TaqMan® OpenArray® Human Signal Transduction Panel was developed to identify differentially expressed genes involved in major signaling pathways. This gene signature panel contained 573 TaqMan® assays specific to signal transduction-related genes plus 24 endogenous controls. The format of the OpenArray® plate allowed for 1 to 4 replicates to be run in parallel per plate. In this real-time PCR assay panel, genes encoding the JAK-STAT, NFκB, Akt, GPCR, cAMP, and MAP kinase pathways were well represented. The full lengths of these pathways were covered, from ligand to receptor to kinase to transcription factor. In addition, there are 18 endogenous control genes against which the assays can be normalized.

The TaqMan® OpenArray® Human Inflammation Panel is designed for quantitative gene expression analysis of inflammation genes important in drug discovery. The panel covered 586 genes that have been studied as targets for a range of inflammatory diseases, plus 21 endogenous control genes, which the assays can be normalized.

4.4.2.2 PCR array analysis

Total RNA was extracted from the transfected colorectal cells using the RNeasy® mini kit (Qiagen, Germany). Harvested total RNA was quantitated using the Nanodrop 100 spectrophotometer (ThermoScientific, Waltham, USA). Reverse transcription was performed using Super-Script

VILO cDNA synthesis kit (Invitrogen, USA) in accordance to the manufacturer's instructions. For PCR array, Human Signal Transduction Panel and Human Inflammation Panel (Life Technologies, USA) were used and PCR was performed on the QuantStudio™ 12K Flex Real-Time PCR System (Life Technologies, USA), according to the manufacturer's instruction (Figure 4.1). Analyses of the raw data were done through the Gene Set Enrichment Analysis Web Portal (Broad Institute of MIT and Harvard) and through the use of Ingenuity Pathway Analysis (IPA) (Ingenuity® Systems, www.ingenuity.com).

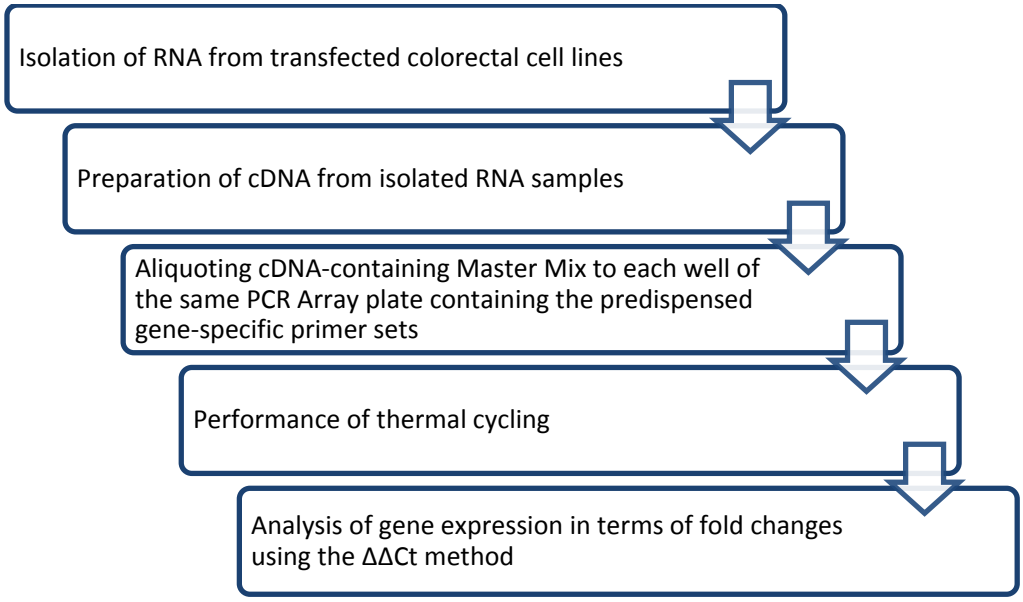


Figure 4.1. The operating procedures of a PCR array

4.4.3 Western Blot Analysis

Cells cultivated to 80% confluence were lysed with M-PER Mammalian Protein Extraction Reagent (Thermo Scientific, USA) containing protease inhibitors (Complete Mini, Roche). Protein concentrations were estimated using a Bradford assay (Bio-Rad, CA, USA) and optical density (OD) was measured with UV-VIS double beam spectrophotometer (Shimadzu, Japan). Aliquots of extracted protein were preserved at −80°C until further use. Samples were mixed with 4x loading buffer (Amresco, Solon, USA), denatured at 99°C for 5 minutes and loaded onto Any kD™ Mini-PROTEAN® TGX™ Precast Gel (BioRad) before electrophoresis. An amount of 40 µg/lane/protein lysate was run. The proteins were transferred to a methanol-activated polyvinylidenedifluoride (PVDF) membrane using an iBlot semi-dry blotting system (Invitrogen, USA) at 10 V for 7 minutes. The membrane was subsequently blocked, probed and washed with reagents provided by the WesternBreeze® kit (Invitrogen, USA). Reagents were used as recommended by the manufacturer.

Primary antibodies (anti-phospho p44/42 MAP Kinase, anti-phospho-SAPK/JNK, anti-phospho- p38 MAPK; 1:1000, Cell Signaling Technology, USA) and secondary antibodies (anti-rabbit HRP-linked IgG, 1:1000, Cell Signaling Technology, USA) were used. Mouse α-tubulin antibody (1:1000; Santa Cruz Biotechnology, USA) was used as a housekeeping protein control. Detection was performed using the Super Signal West Pico Chemiluminescent Substrate kit (Pierce, Thermo Scientific, USA) and chemiluminescent signals were captured by the C-DiGit Blot scanner (Li-Cor Biosciences, USA).

Membranes were stripped using a Restore PLUS Western Blotting Stripping buffer (Pierce, Thermo Scientific, USA) at room temperature for 30 minutes and reprobed with another antibody immediately. All subsequent reprobing procedures started with a blocking step as described above.

4.5 Results

4.5.1 PCR array analysis of colorectal cancer cells expressing HERV-H using GSEA

The gene expression profiles of colorectal cancer cells transfected with (i) an empty vector (pEGFP-N2) and (ii) a vector containing HERV-H (pEGFP-N2:HERV-H) were analysed using a high-throughput PCR array containing 573 primer/probe sets for signal transduction-related genes and 586 primer/probe sets for inflammation-associated genes. These were mapped into the Gene Set Enrichment Analysis (GSEA) software and were filtered according to the manufacturer's instructions. Specifically, bioinformatics analysis revealed the differentially regulated genes were associated with host cellular pathways that are involved in cell cycle/mitosis/proliferation/apoptosis.

The top 20 upregulated (>1.5 fold) signalling transduction pathways identified from 27 signal transduction-related genes in LS174T:HERV-H_GFP were shown in Table 4.2. The upregulated pathways included those involved in cytokine signalling, β-catenin signalling, Wnt-mediated signalling,

G-protein α signalling, PI3K signalling, p53 downstream effectors signalling, Stem cell factor receptor-mediated signalling and MAPK signalling.

Among the genes that were involved in these top 20 identified pathways were gamma interleukin 2 receptor (IL2RG), β interleukin 1 (IL1B), wingless-type MMTV integration site family member 11 (Wnt11), MAP4K1 and nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 4 (NFATC4) (Figure 4.2).

The heat map of differentially expressed genes using LS174T:HERV-H_GFP and LS174T:GFP cells is shown in Figure 4.3. Heat map analysis indicated that early growth response 1 (EGR1), dickkopf WNT signaling pathway inhibitor 1 (DKK1), caspase 1 (CASP1), interleukin 1 β (IL1B), NFATC4 and Wnt11 were differentially upregulated whereas myocyte enhancer factor 2C (MEF2C), paired-like homeodomain transcription factor 2 (PITX2) and mitogen-activated protein kinase kinase kinase kinase 1 (MAP4K1) were differentially downregulated.

Table 4.2. Top 20 signal transduction pathways. The pathways were identified from 27 genes with ΔCT>1.5 fold in LS174T:HERV-H_GFP. FDR: false discovery rate q value set at 0.25; *P* value set at 0.05

Gene Set Name [# Genes (K)]	Description	# Genes in Overlap (k)	k/K	p-value	FDR q-value
REACTOME_IMMUNE_SYSTEM [933]	Genes involved in Immune System	9		$1.9\ e^{-9}$	$2.51\ e^{-6}$
REACTOME_CYTOKINE_SIGNALING_IN_IMMUNE_NE_SYSTEM [270]	Genes involved in Cytokine Signaling in Immune system	6		$1.04\ e^{-8}$	$6.85\ e^{-6}$
REACTOME_GPCR_LIGAND_BINDING [408]	Genes involved in GPCR ligand binding	6		$1.19\ e^{-7}$	$4.78\ e^{-5}$
PID_BETACATENIN_NUC_PATHWAY [80]	Regulation of nuclear beta catenin signaling and target gene transcription	4		$1.45\ e^{-7}$	$4.78\ e^{-5}$
WNT_SIGNALING [89]	Genes related to Wnt-mediated signal transduction	4		$2.23\ e^{-7}$	$5.88\ e^{-5}$
KEGG_NEUROACTIVE_LIGAND_RECEPTOR_INTERTERACTION [272]	Neuroactive ligand-receptor interaction	5		$5.08\ e^{-7}$	$1.12\ e^{-4}$
REACTOME_G_ALPHA_S_SIGNALLING_EVENTS [121]	Genes involved in G alpha (s) signalling events	4		$7.66\ e^{-7}$	$1.44\ e^{-4}$
PID_IL2_PI3KPATHWAY [34]	IL2 signaling events mediated by PI3K	3		$1.07\ e^{-6}$	$1.76\ e^{-4}$
PID_P53DOWNSTREAMPATHWAY [137]	Direct p53 effectors	4		$1.26\ e^{-6}$	$1.84\ e^{-4}$
PID_KITPATHWAY [52]	Signaling events mediated by Stem cell factor receptor (c-Kit)	3		$3.92\ e^{-6}$	$5.17\ e^{-4}$
PID_IL12_2PATHWAY [63]	IL12-mediated signaling events	3		$7.01\ e^{-6}$	$7.72\ e^{-4}$
PID_MYC_REPRESSPATHWAY [63]	Validated targets of C-MYC transcriptional repression	3		$7.01\ e^{-6}$	$7.72\ e^{-4}$
PID_CD8TCRDOWNSTREAMPATHWAY [65]	Downstream signaling in naïve CD8+ T cells	3		$7.71\ e^{-6}$	$7.83\ e^{-4}$
REACTOME_SIGNALING_BY_GPCR [920]	Genes involved in Signaling by GPCR	6		$1.31\ e^{-5}$	$1.23\ e^{-3}$
PID_SMAD2_3NUCLEARPATHWAY [82]	Regulation of nuclear SMAD2/3 signaling	3		$1.55\ e^{-5}$	$1.37\ e^{-3}$
KEGG_MAPK_SIGNALING_PATHWAY [267]	MAPK signaling pathway	4		$1.76\ e^{-5}$	$1.39\ e^{-3}$
KEGG_TGF_BETA_SIGNALING_PATHWAY [86]	TGF-beta signaling pathway	3		$1.79\ e^{-5}$	$1.39\ e^{-3}$
KEGG_GAP_JUNCTION [90]	Gap junction	3		$2.05\ e^{-5}$	$1.51\ e^{-3}$
REACTOME_SIGNALING_BY_ILS [107]	Genes involved in Signaling by Interleukins	3		$3.45\ e^{-5}$	$2.39\ e^{-3}$
BIOCARTA_CK1_PATHWAY [17]	Regulation of ck1/cdk5 by type 1 glutamate receptors	2		$4.5\ e^{-5}$	$2.58\ e^{-3}$

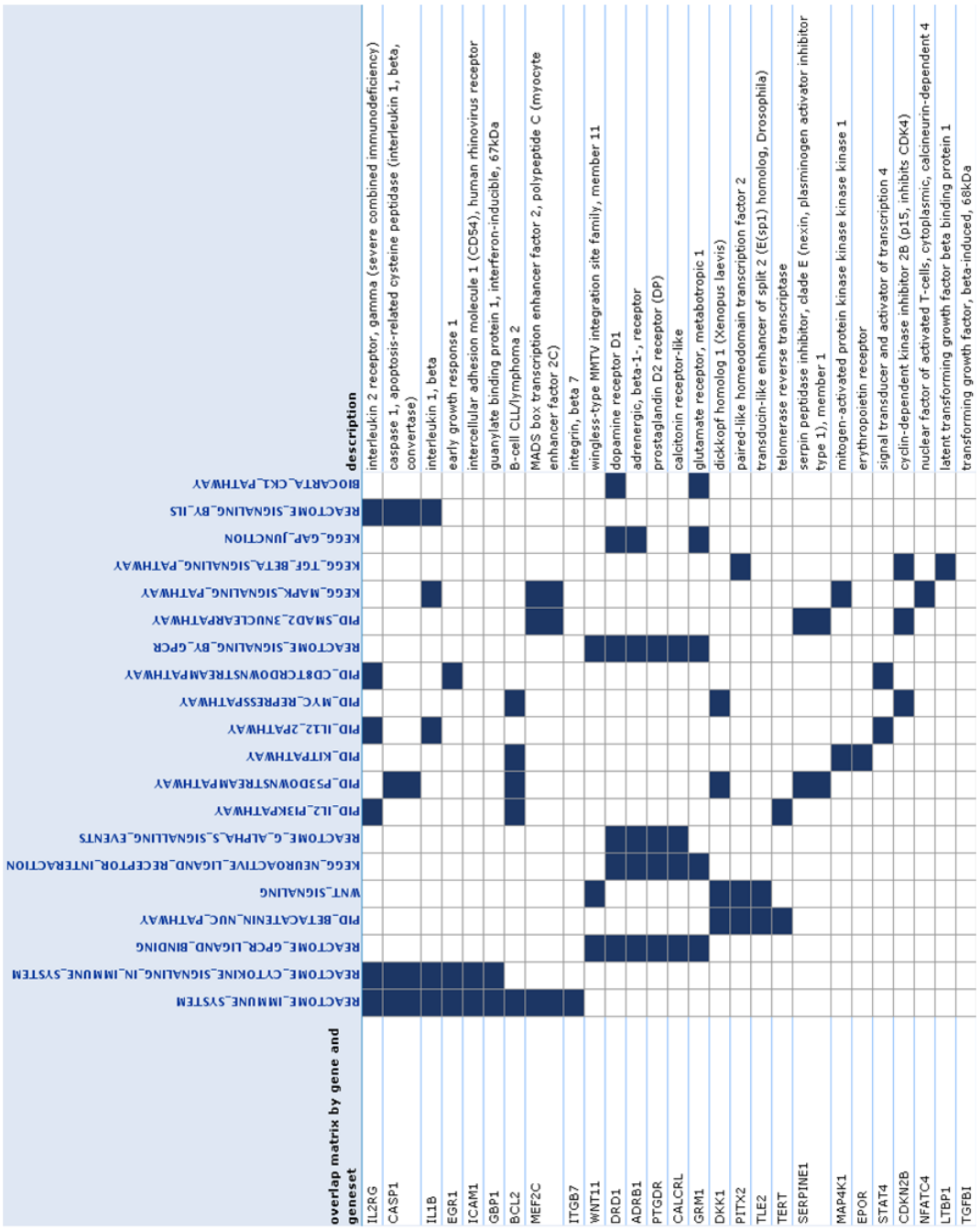


Figure 4.2. The distribution of each signal transduction-related gene in the top 20 identified pathways.

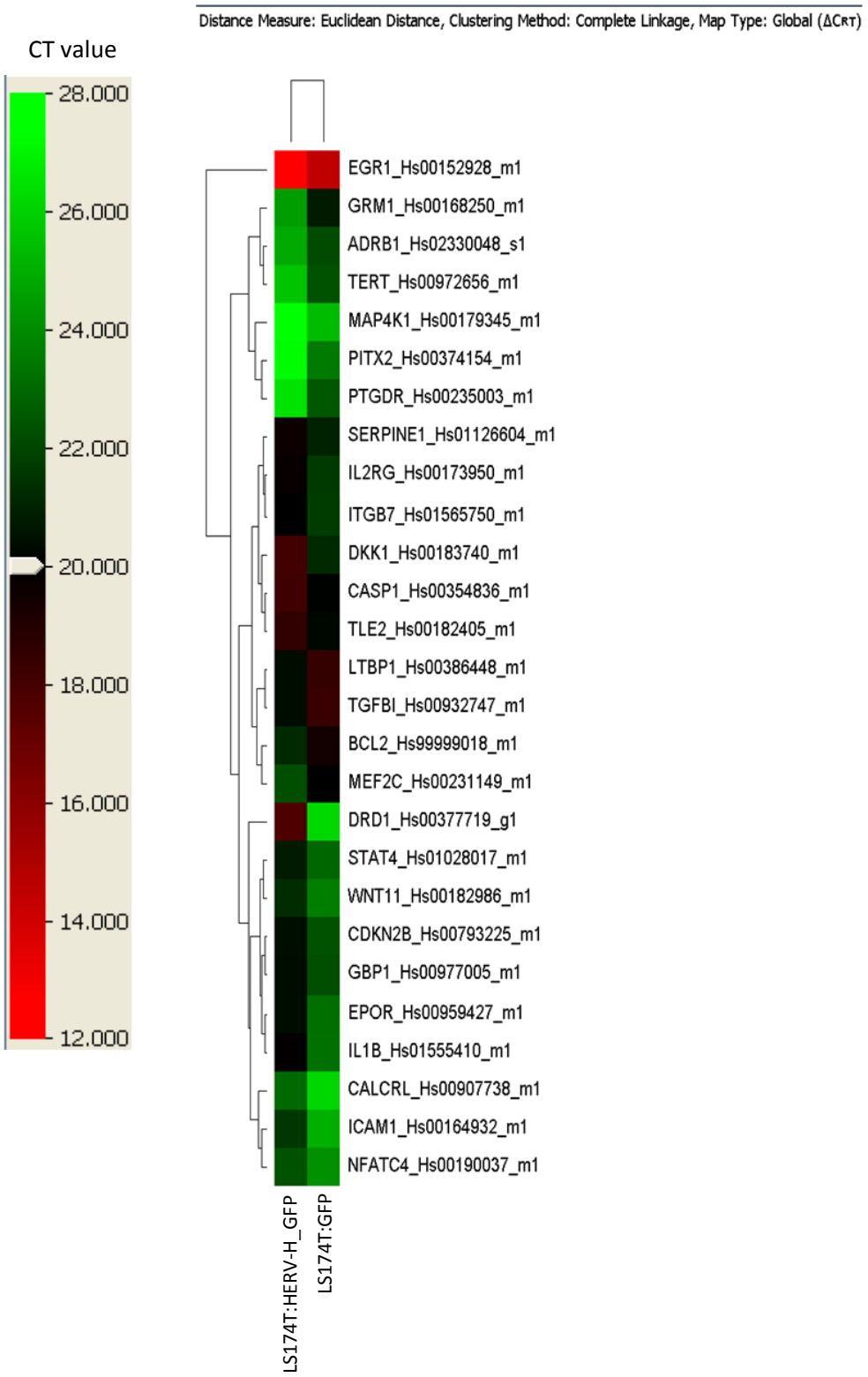


Figure 4.3. Differential expression of signal transduction-related genes in transfected LS174T colorectal cancer cells. Heat map, and gene expression increased (red) and decreased (green) in a comparison of colorectal cells expressing HERV-H versus cells not expressing HERV-H.

The top 20 upregulated (>1.5 fold) inflammatory diseases associated pathways identified from 28 inflammation-related genes in LS174T:HERV-H_GFP are shown in Table 4.3.

Among the upregulated pathways were the interleukin 2-mediated signalling pathway, the interleukin 1-mediated signalling pathway, the JAK-STAT signalling pathway, the interleukin 10 anti-inflammatory signalling pathway and the Wnt signalling pathway.

Among the genes that were involved in the top 20 identified pathways were β interleukin 1 (IL1B), type I interleukin 1 receptor (IL1R1), v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog (KIT), TP53, NFATC4 and Wnt 16 (Figure 4.4).

Heat map analysis indicated that NFATC4, interleukin 16 (IL16), caspase 1 (CASP1) and interleukin 1 β (IL1B) were differentially upregulated whereas cell death-inducing p53 target 1 (C16orf5), tumour protein p53 (TP53) and KIT were differentially downregulated (Figure 4.5).

Table 4.3. Top 20 inflammation associated pathways. The pathways were identified from 28 genes with ΔCT>1.5 fold in LS174T:HERV-H_GFP. FDR: false discovery rate q value set at 0.25; *P* value set at 0.05

Gene Set Name [# Genes (K)]	Description	# Genes in Overlap (k)	k/K	p-value	FDR q-value
KEGG_CYTOKINE_CYTOKINE_RECEPTOR_INTERACTION [267]	Cytokine-cytokine receptor interaction	10		$3.33\ e^{-16}$	$4.4\ e^{-13}$
PID_IL12_2PATHWAY [63]	IL12-mediated signaling events	4		$5.49\ e^{-8}$	$3.63\ e^{-5}$
KEGG_HEMATOPOIETIC_CELL_LINEAGE [88]	Hematopoietic cell lineage	4		$2.13\ e^{-7}$	$9.37\ e^{-5}$
REACTOME_SIGNALING_BY_ILS [107]	Genes involved in Signaling by Interleukins	4		$4.68\ e^{-7}$	$1.54\ e^{-4}$
PID_IL1PATHWAY [34]	IL1-mediated signaling events	3		$1.07\ e^{-6}$	$2.82\ e^{-4}$
KEGG_JAK_STAT_SIGNALING_PATHWAY [155]	Jak-STAT signaling pathway	4		$2.06\ e^{-6}$	$4.52\ e^{-4}$
KEGG_MAPK_SIGNALING_PATHWAY [267]	MAPK signaling pathway	4		$1.76\ e^{-5}$	$2.82\ e^{-3}$
REACTOME_CYTOKINE_SIGNALING_IN_IMMUNE_NE_SYSTEM [270]	Genes involved in Cytokine Signaling in Immune system	4		$1.84\ e^{-5}$	$2.82\ e^{-3}$
KEGG_APOPTOSIS [88]	Apoptosis	3		$1.92\ e^{-5}$	$2.82\ e^{-3}$
KEGG_PATHWAYS_IN_CANCER [328]	Pathways in cancer	4		$3.93\ e^{-5}$	$4.57\ e^{-3}$
BIOCARTA_IL10_PATHWAY [17]	IL-10 Anti-inflammatory Signaling Pathway	2		$4.5\ e^{-5}$	$4.57\ e^{-3}$
BIOCARTA_NO2IL12_PATHWAY [17]	NO2-dependent IL 12 Pathway in NK cells	2		$4.5\ e^{-5}$	$4.57\ e^{-3}$
PID_ANTHRAXPATHWAY [17]	Cellular roles of Anthrax toxin	2		$4.5\ e^{-5}$	$4.57\ e^{-3}$
BIOCARTA_CYTOKINE_PATHWAY [22]	Cytokine Network	2		$7.62\ e^{-5}$	$7.19\ e^{-3}$
BIOCARTA_IL12_PATHWAY [23]	IL12 and Stat4 Dependent Signaling Pathway in Th1 Development	2		$8.35\ e^{-5}$	$7.34\ e^{-3}$
KEGG_WNT_SIGNALING_PATHWAY [151]	Wnt signaling pathway	3		$9.6\ e^{-5}$	$7.92\ e^{-3}$
PID_IL27PATHWAY [26]	IL27-mediated signaling events	2		$1.07\ e^{-4}$	$8.32\ e^{-3}$
PID_IL2_STAT5PATHWAY [30]	IL2 signaling events mediated by STAT5	2		$1.43\ e^{-4}$	$1.05\ e^{-2}$
BIOCARTA_IL1R_PATHWAY [33]	Signal transduction through IL1R	2		$1.74\ e^{-4}$	$1.21\ e^{-2}$
REACTOME_IMMUNE_SYSTEM [933]	Genes involved in Immune System	5		$1.9\ e^{-4}$	$1.23\ e^{-2}$

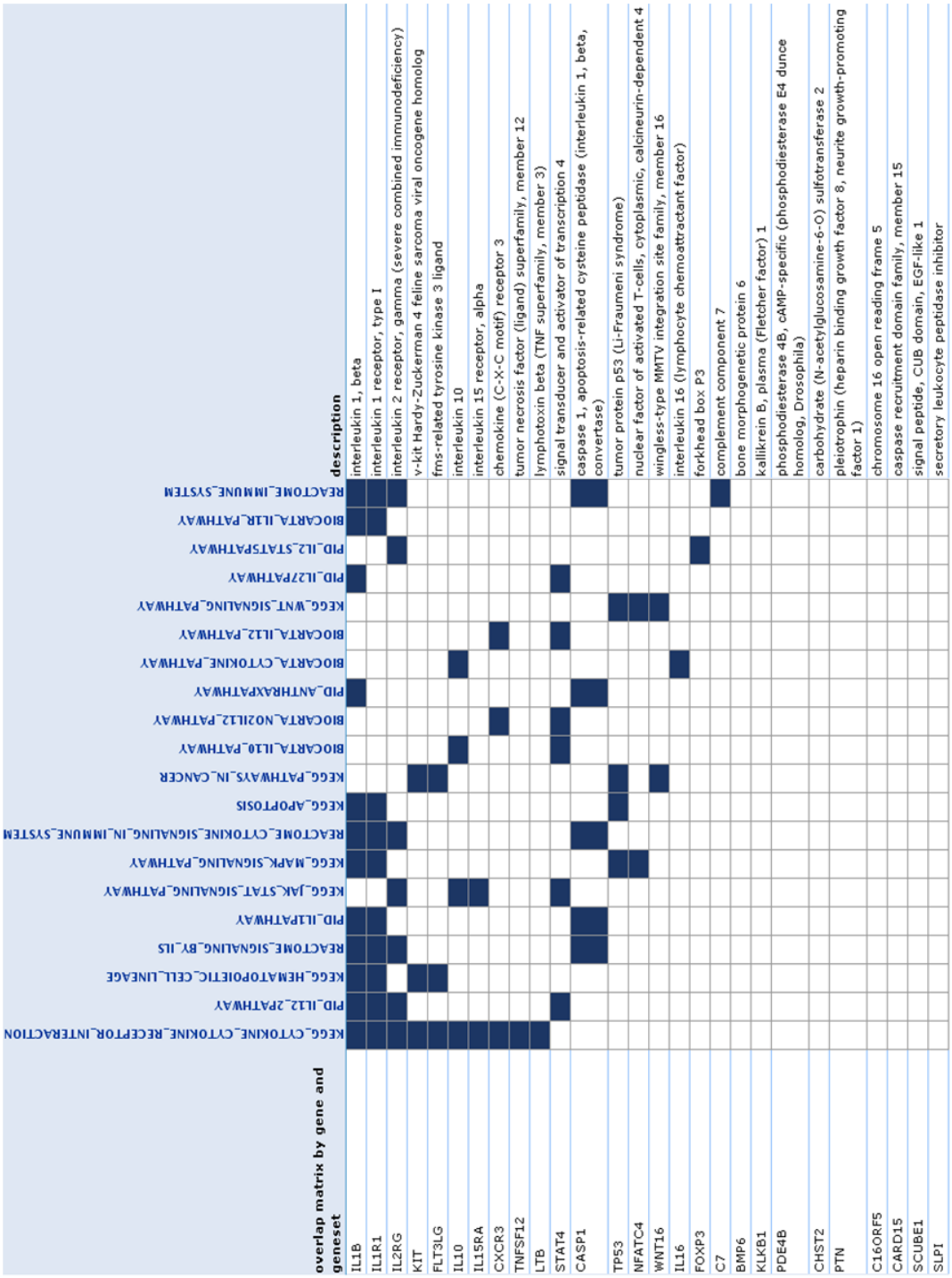


Figure 4.4. The distribution of each inflammatory diseases-related gene in the top 20 identified pathways.

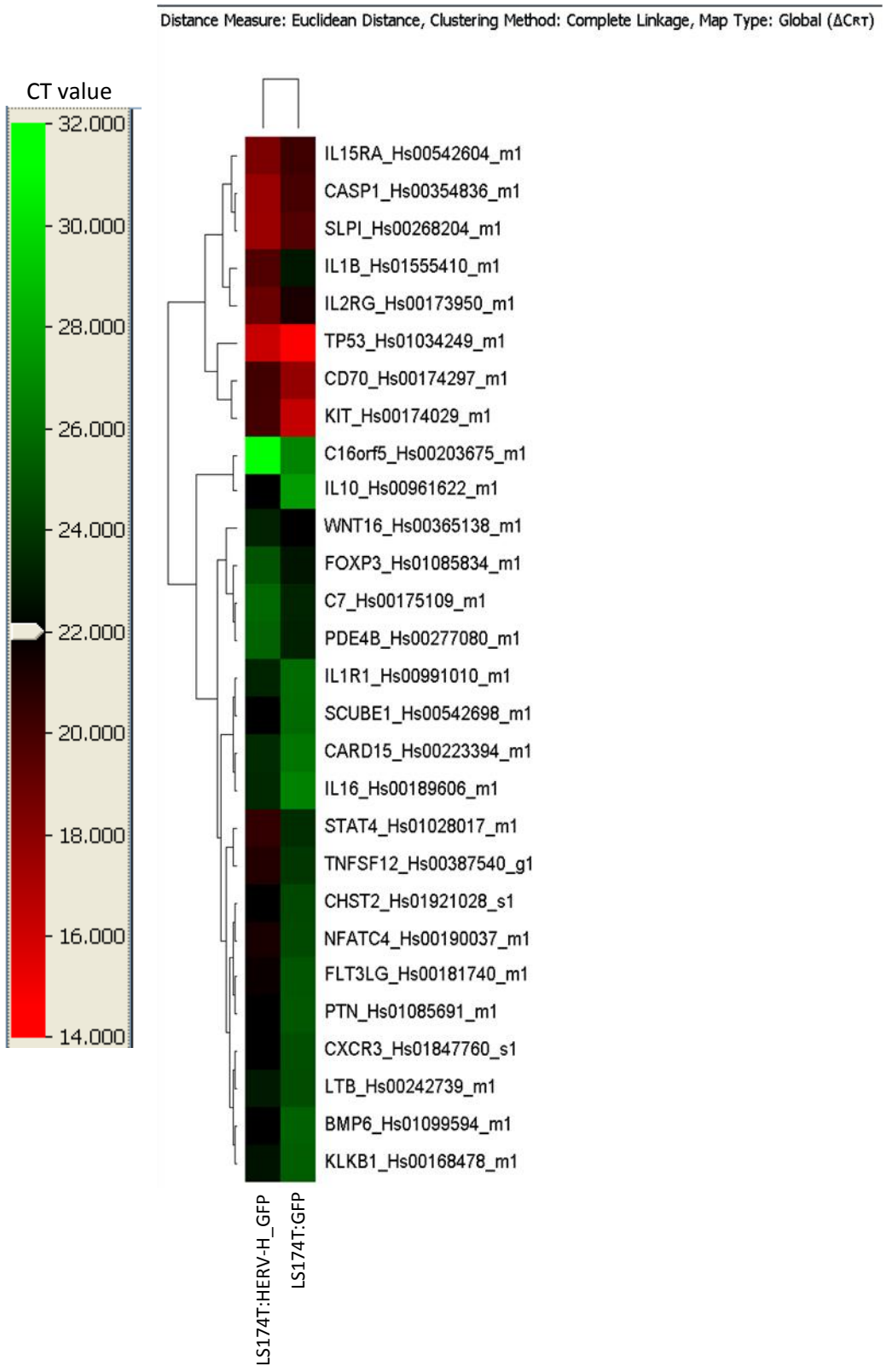


Figure 4.5. Differential expression of inflammatory diseases associated genes in transfected LS174T colorectal cancer cells. Heat map, and gene expression increased (red) and decreased (green) in a comparison of colorectal cells expressing HERV-H versus cells not expressing HERV-H.

4.5.2 PCR array analysis of colorectal cancer cells expressing HERV-H using IPA

Using the ingenuity pathway analysis (IPA) application, the cellular pathways that were affected by HERV-H mediated transforming process were identified. With the use of Human Signal Transduction Panel, the top scoring network (Score = 31) associated with HERV-H transduction include genes involved in cell death and survival, inflammatory response and cancer (Table 4.4). In addition, cancer was the top disease implicated. There were four top molecular and cellular functions significantly associated with HERV-H transduction. These four functions were associated with cell death and survival, cellular growth and proliferation, cellular development and cell cycle (Table 4.4).

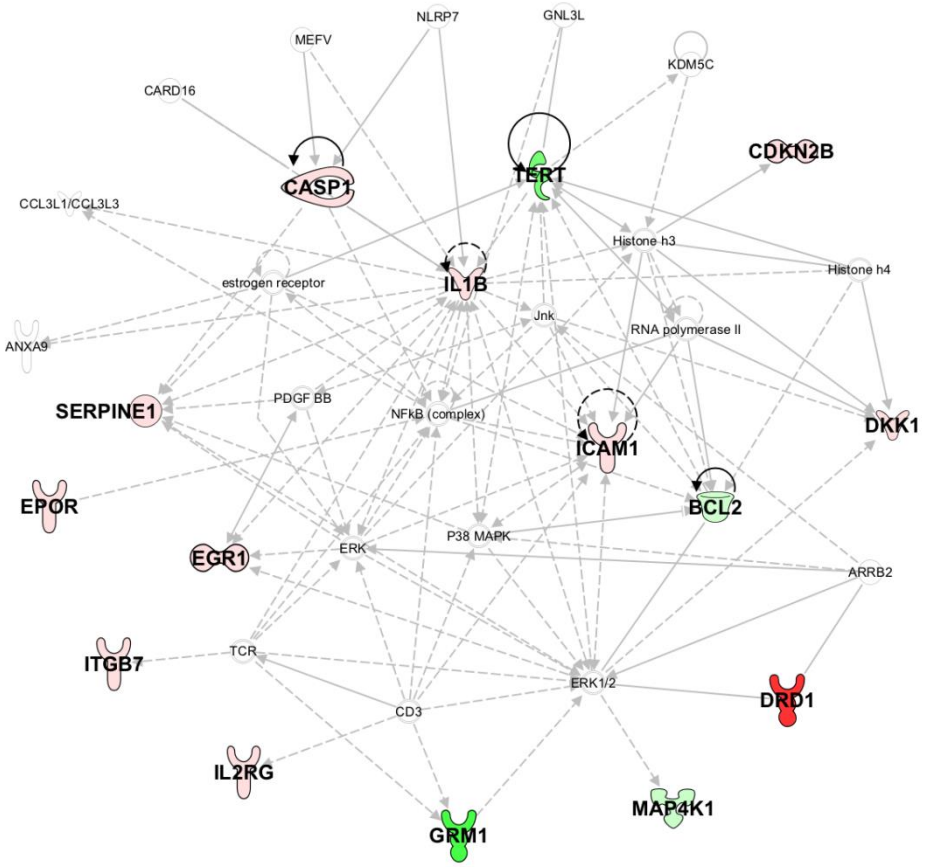
With the canonical pathway analysis, the top pathway that was identified to be significantly associated with HERV-H transduction was the TGF-β signalling in colorectal cancer cells (Table 4.4). Signalling network showing various canonical pathways following HERV-H transduction is presented in Figure 4.6.

Table 4.4. IPA analysis of implicated network/pathway in human signal transduction panel after HERV-H transduction.

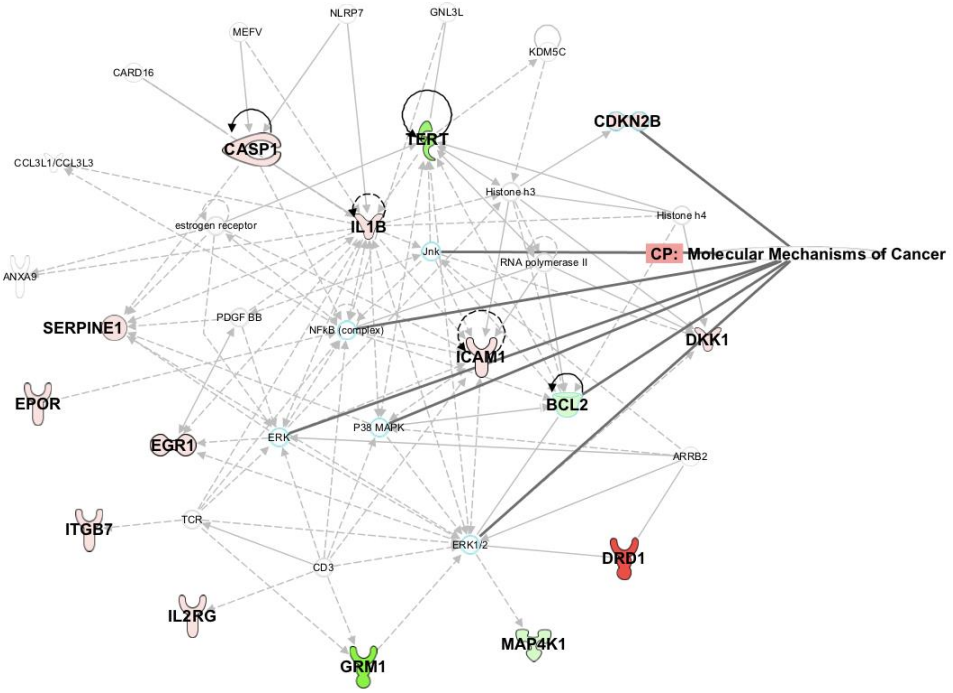
Associated Network		**Score**
Cell death and survival, Inflammatory response, Cancer		31
Diseases and Disorders	***P* value**	**Molecules involved**
Cancer	1.32×10^{-07} - 6.59×10^{-03}	25
Molecular and Cellular Functions	***P* value**	**Molecules involved**
Cell Death and Survival	6.22×10^{-07} - 6.60×10^{-03}	15
Cellular Growth and Proliferation	8.66×10^{-07} - 4.98×10^{-03}	15
Cellular Development	2.43×10^{-06} - 5.13×10^{-03}	15
Cell Cycle	2.63×10^{-06} - 4.95×10^{-03}	5
Top Canonical Pathways	***P* value**	**Ratio**
TGF-β signalling	1.2×10^{-05}	4/93 (0.043)
Role of Osteoblasts, Osteoclasts and Chondrocytes in Rheumatoid Arthritis	2.4×10^{-05}	5/225 (0.022)

IPA employs Fisher's exact test to determine the relationship between the input dataset and the canonical pathways with associated biological functions.

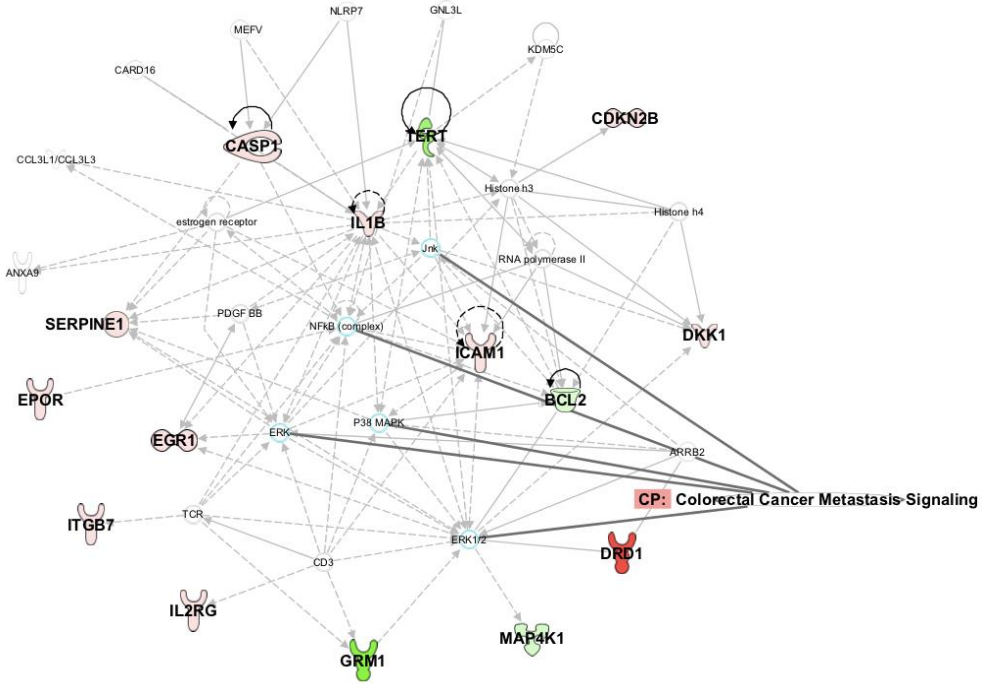
(A)



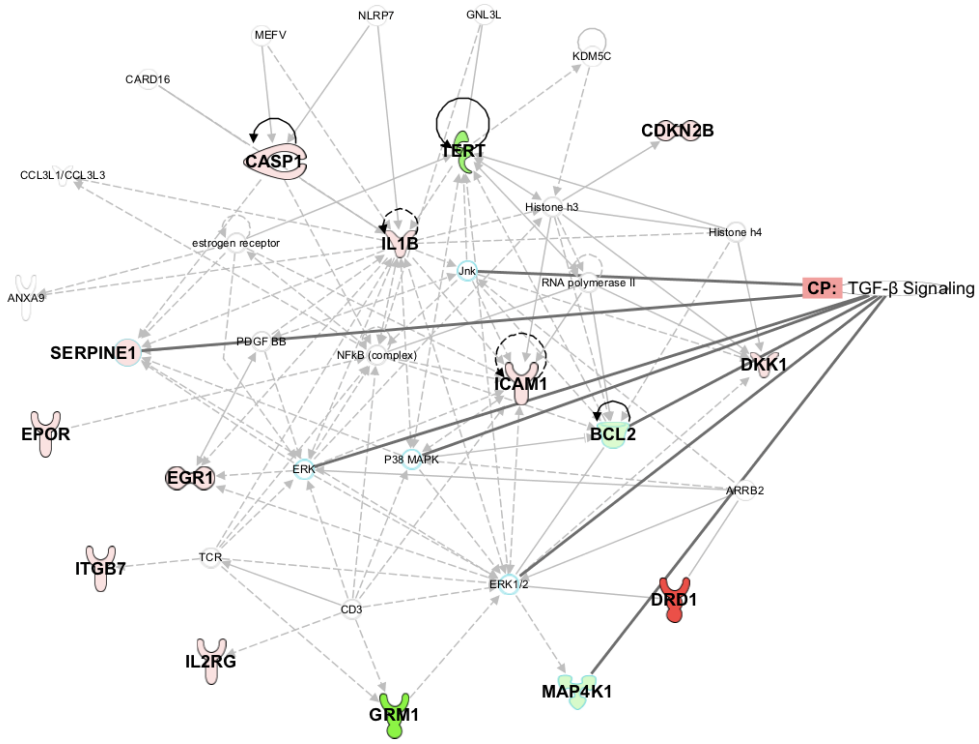
(B)


NLRP7
GNL3L
MEFV
KDM5C
CARD16
CASP1
TERT
CDKN2B
CCL3L1/CCL3L3
Histone h3
Histone h4
estrogen receptor
IL1B
Jnk
RNA polymerase II
CP: Molecular Mechanisms of Cancer
ANXA9
SERPINE1
PDGF BB
NFkB (complex)
ICAM1
DKK1
EPOR
BCL2
P38 MAPK
EGR1
ERK
ARRB2
ITGB7
TCR
ERK1/2
DRD1
CD3
IL2RG
MAP4K1
GRM1

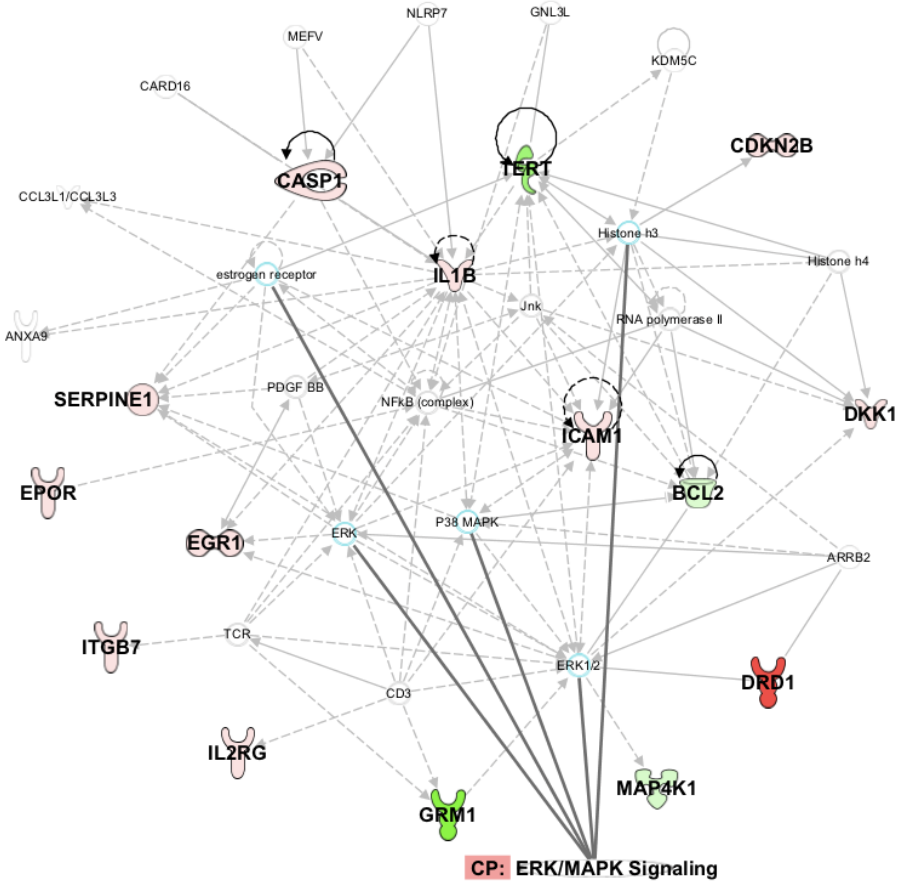
(C)


NLRP7
GNL3L
MEFV
KDM5C
CARD16
CASP1
TERT
CDKN2B
CCL3L1/CCL3L3
Histone h3
Histone h4
estrogen receptor
IL1B
Jnk
RNA polymerase II
ANXA9
SERPINE1
PDGF BB
NFkB (complex)
ICAM1
DKK1
EPOR
BCL2
P38 MAPK
EGR1
ERK
ARRB2
CP: Colorectal Cancer Metastasis Signaling
ITGB7
TCR
ERK1/2
DRD1
CD3
IL2RG
MAP4K1
GRM1

(D)


NLRP7
GNL3L
MEFV
KDM5C
CARD16
CDKN2B
CASP1
TERT
CCL3L1/CCL3L3
Histone h3
estrogen receptor
IL1B
Histone h4
Jnk
RNA polymerase II
CP: TGF-β Signaling
ANXA9
SERPINE1
PDGF BB
NFkB (complex)
ICAM1
DKK1
EPOR
BCL2
P38 MAPK
EGR1
ERK
ARRB2
ITGB7
TCR
ERK1/2
DRD1
CD3
IL2RG
MAP4K1
GRM1

(E)


NLRP7
GNL3L
MEFV
KDM5C
CARD16
CDKN2B
CASP1
TERT
CCL3L1/CCL3L3
Histone h3
estrogen receptor
IL1B
Histone h4
Jnk
RNA polymerase II
ANXA9
SERPINE1
PDGF BB
NFkB (complex)
ICAM1
DKK1
EPOR
BCL2
P38 MAPK
EGR1
ERK
ARRB2
ITGB7
TCR
ERK1/2
DRD1
CD3
IL2RG
MAP4K1
GRM1
CP: ERK/MAPK Signaling

(F)

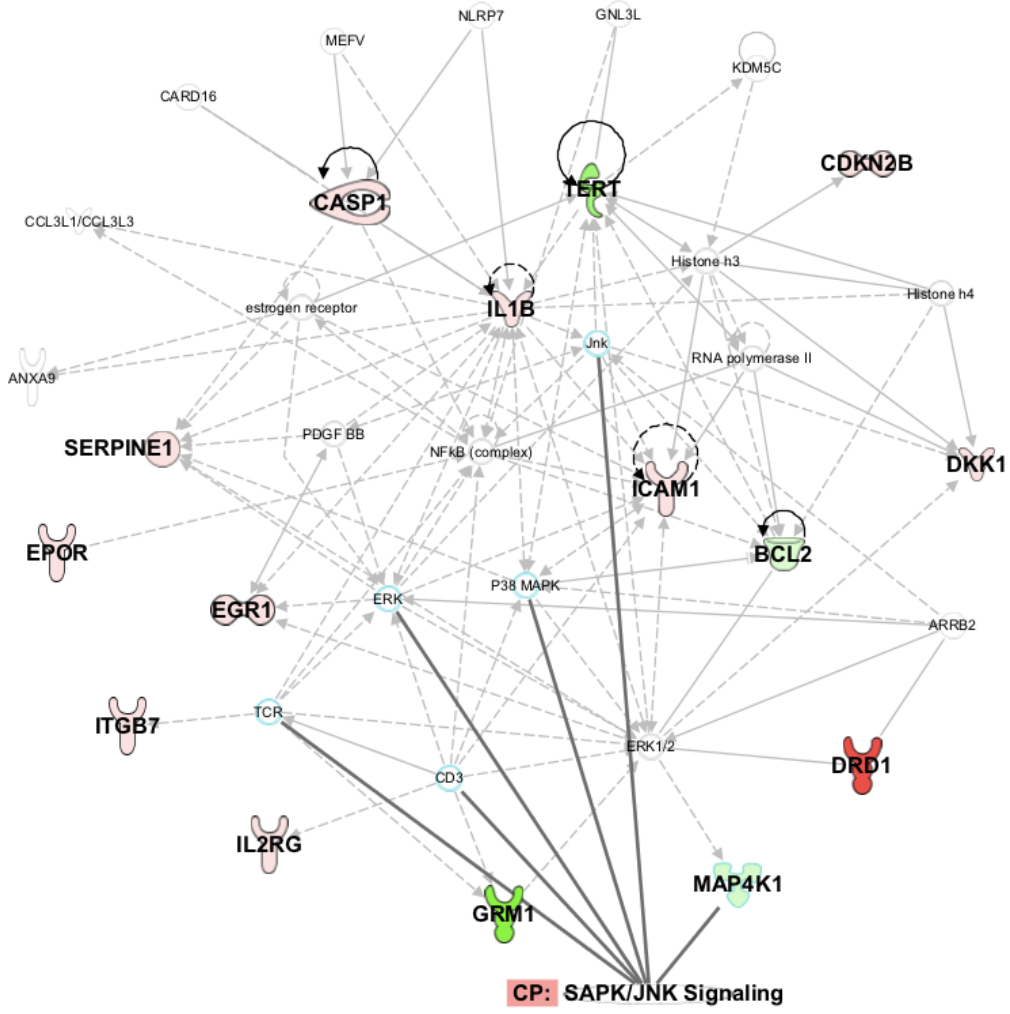


Figure 4.6. Connectivity of differentially expressed signal transduction-related genes in colorectal cancer cells following HERV-H transduction. (A) Signalling network showing all involved genes. (B) Signalling network showing canonical pathway of molecular mechanisms of cancer. (C) Signalling network showing canonical pathway of colorectal cancer metastasis signalling. (D) Signalling network showing canonical pathway of TGF-β signalling. (E) Signalling network showing canonical pathway of ERK/MAPK signalling. (F) Signalling network showing canonical pathway of SAPK/JNK signalling. Geometric figures in bold letters denote genes involvement in the study. Solid interconnecting lines shows the genes that are directly connected and the dotted lines signify indirect connection between the genes and cellular functions.

On the other hand, using the Human Inflammation Panel, the top scoring network (score = 43) associated with HERV-H transduction include genes involved in cellular movement, haematological system development and function and immune cell trafficking (Table 4.5). Moreover, both inflammatory disease and response were the top ranking disease and disorder (Table 4.5).

Cellular development, cellular growth and proliferation, cell-to-cell signalling and interaction, cell death and survival, and cellular movement were the predominant biological functions associated with the top five molecular and cellular functions (Table 4.5).

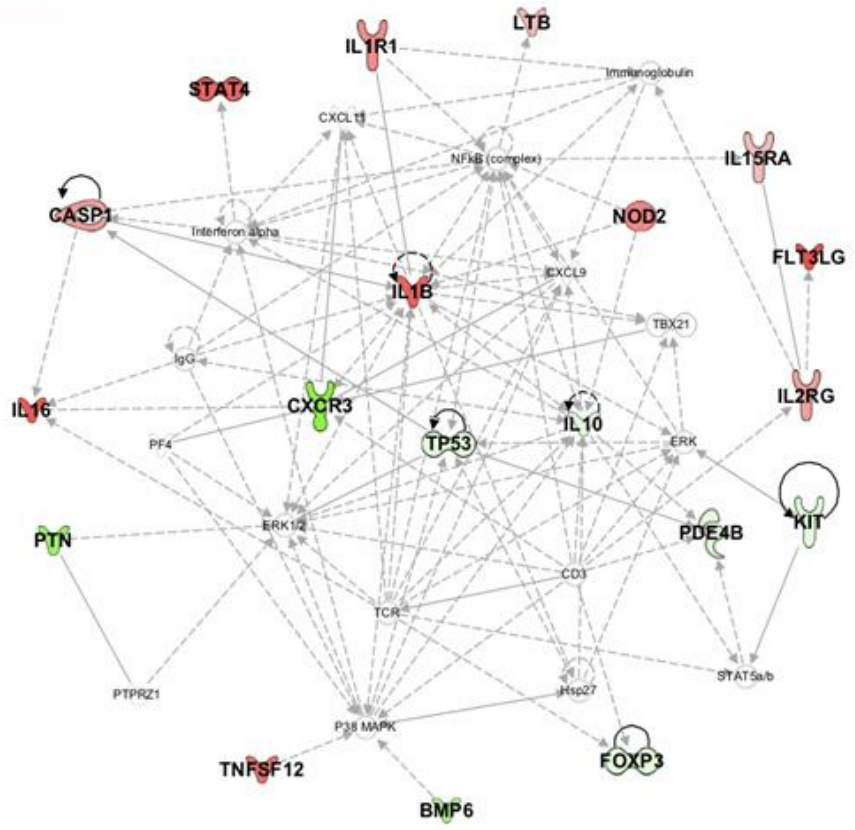
With the canonical pathway analysis, the top pathway that was identified to be significantly associated with HERV-H transduction was intriguingly involving the role of macrophages, fibroblasts and endothelial cells in rheumatoid arthritis (Table 4.5). Signalling network showing various canonical pathways following HERV-H transduction is presented in Figure 4.7.

Table 4.5. IPA analysis of implicated network/pathway in human inflammation panel after HERV-H transduction.

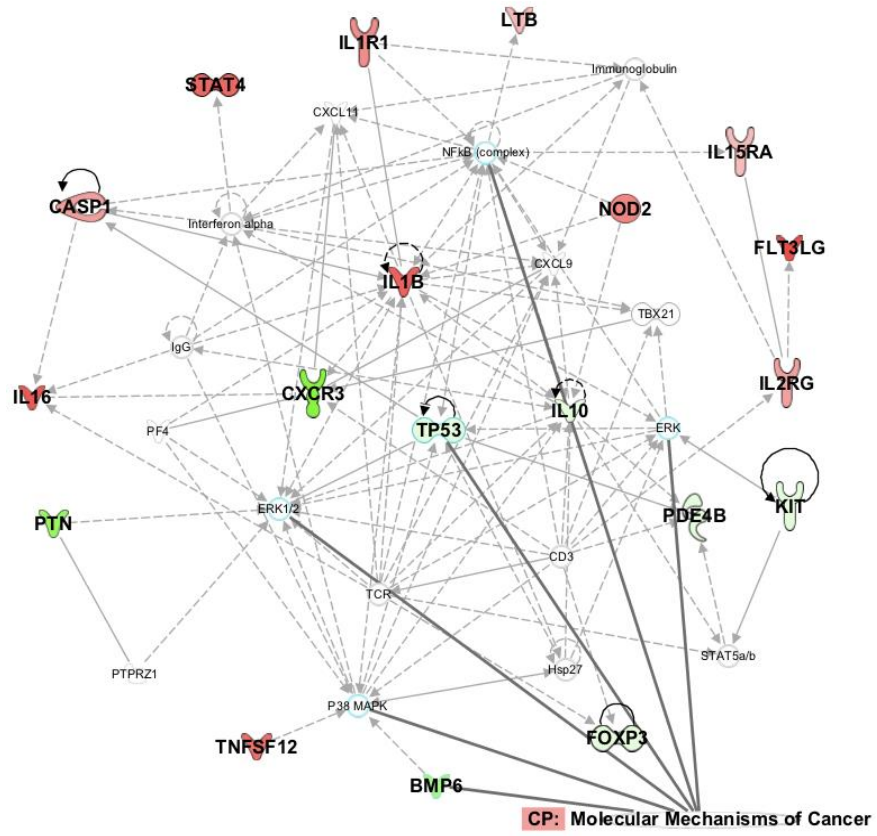
Associated Network		**Score**
Cellular Movement, Haematological System Development and Function, Immune Cell Trafficking		43
Cancer, Cellular Development, Cellular Growth and Proliferation		6
Diseases and Disorders	***P* value**	**Molecules involved**
Inflammatory Disease	9.01×10^{-11} - 3.43×10^{-03}	19
Inflammatory Response	1.13×10^{-07} - 3.43×10^{-03}	17
Molecular and Cellular Functions	***P* value**	**Molecules involved**
Cellular Development	1.43×10^{-09} - 3.43×10^{-03}	18
Cellular Growth and Proliferation	1.43×10^{-09} - 3.43×10^{-03}	18
Cell-To-Cell Signaling and Interaction	5.43×10^{-09} - 3.43×10^{-03}	18
Cell Death and Survival	3.42×10^{-08} - 3.43×10^{-03}	16
Cellular Movement	6.69×10^{-07} - 3.43×10^{-03}	12
Top Canonical Pathways	***P* value**	**Ratio**
Role of Macrophages, Fibroblasts and Endothelial Cells in Rheumatoid Arthritis	4.24×10^{-07}	7/311 (0.023)
Role of Osteoblasts, Osteoclasts and Chondrocytes in Rheumatoid Arthritis	1.43×10^{-06}	6/225 (0.027)
TREM1 Signalling	2.22×10^{-06}	4/57 (0.07)

IPA employs Fisher's exact test to determine the relationship between the input dataset and the canonical pathways with associated biological functions.

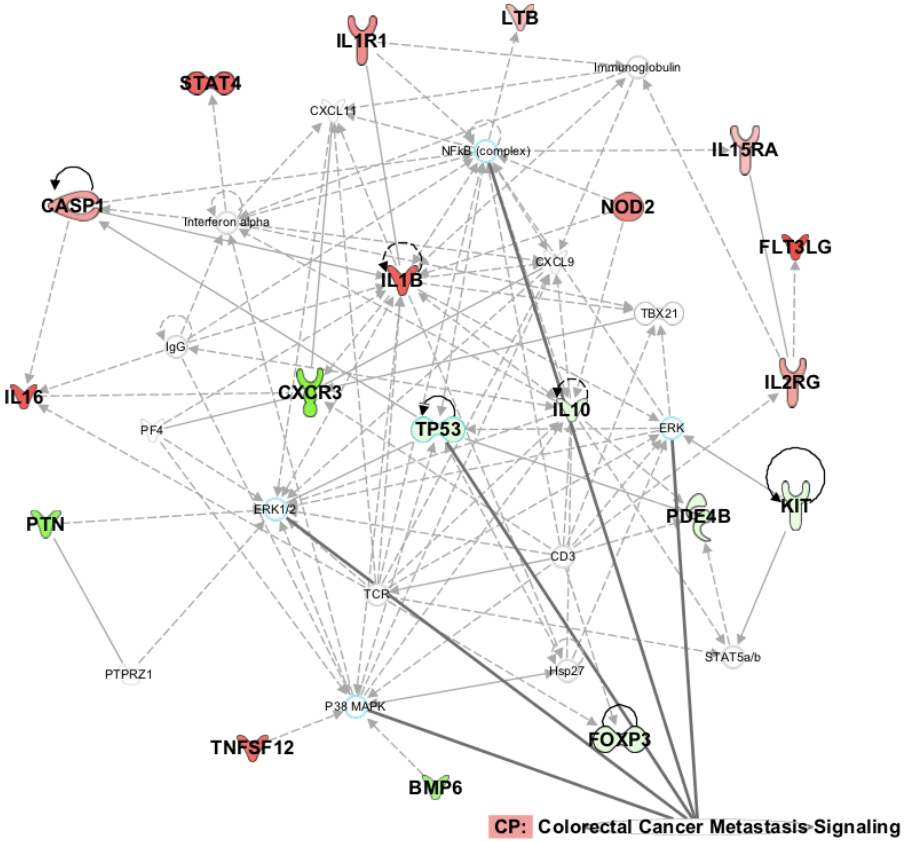
(A)



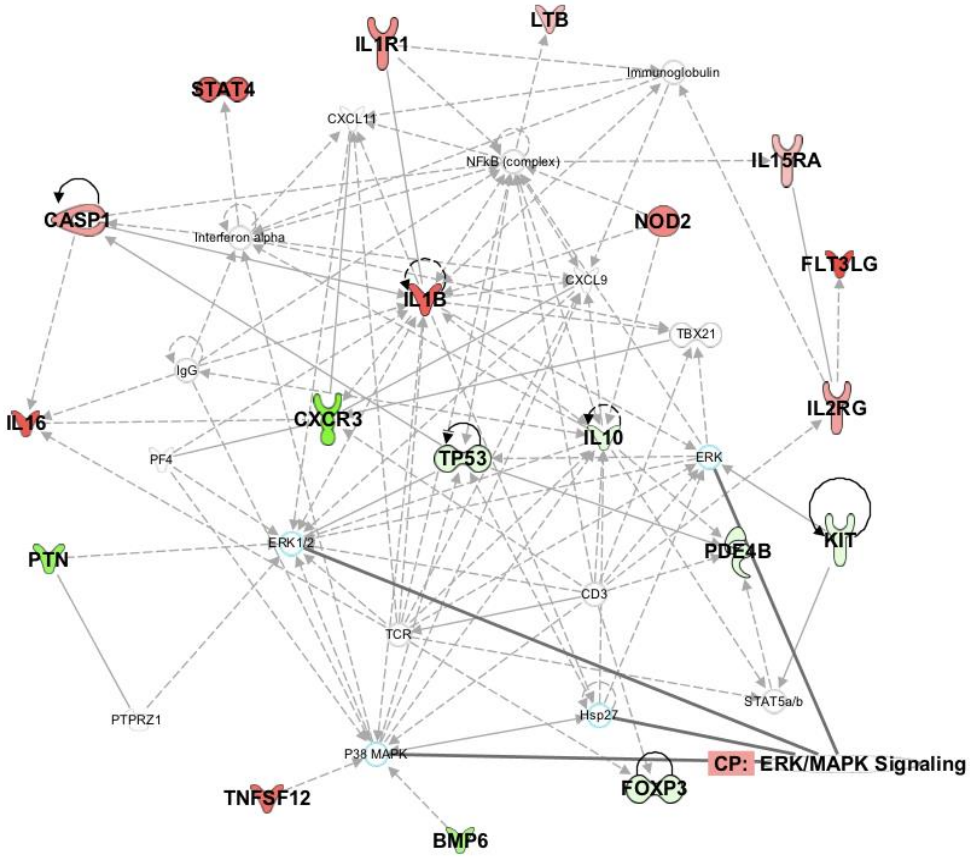
(B)


LTB
IL1R1
Immunoglobulin
STAT4
CXCL11
NFkB (complex)
IL15RA
CASP1
NOD2
Interferon alpha
FLT3LG
CXCL9
IL1B
TBX21
IgG
IL2RG
IL16
CXCR3
IL10
TP53
ERK
PF4
KIT
PTN
ERK1/2
PDE4B
CD3
TCR
STAT5a/b
PTPRZ1
Hsp27
P38 MAPK
FOXP3
TNFSF12
BMP6
CP: Molecular Mechanisms of Cancer

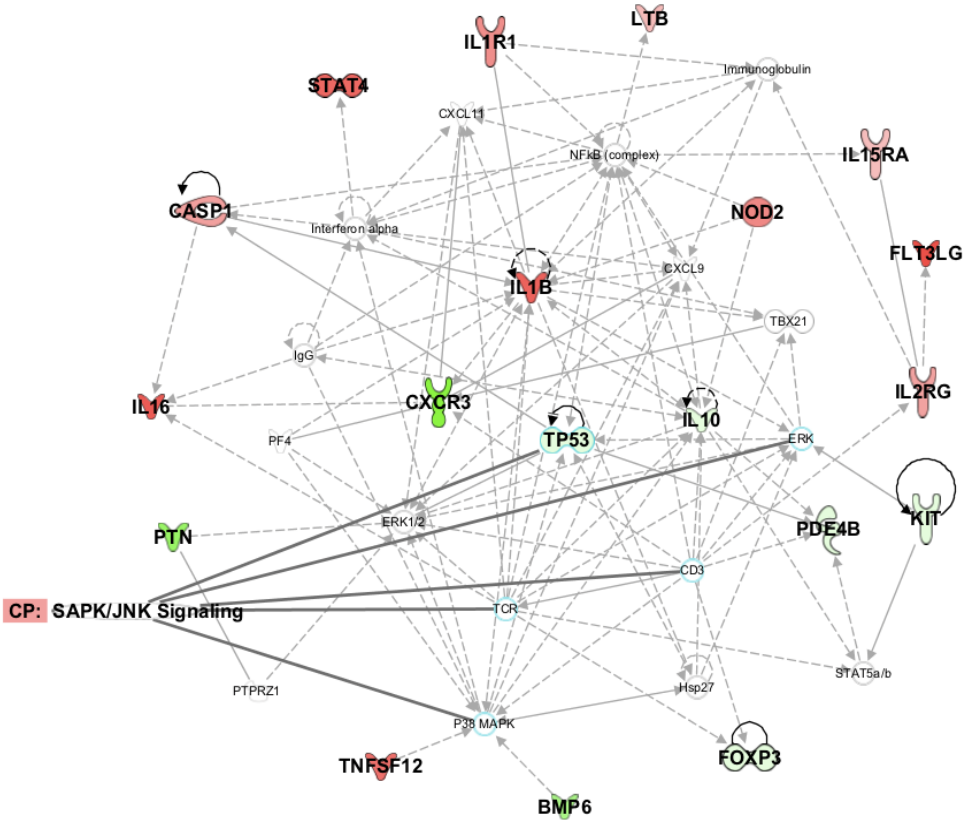
(C)


LTB
IL1R1
Immunoglobulin
STAT4
CXCL11
NFkB (complex)
IL15RA
CASP1
NOD2
Interferon alpha
FLT3LG
CXCL9
IL1B
TBX21
IgG
IL2RG
IL16
CXCR3
IL10
TP53
ERK
PF4
KIT
PTN
ERK1/2
PDE4B
CD3
TCR
STAT5a/b
PTPRZ1
Hsp27
P38 MAPK
FOXP3
TNFSF12
BMP6
CP: Colorectal Cancer Metastasis Signaling

(D)


LTB
IL1R1
Immunoglobulin
STAT4
CXCL11
NFkB (complex)
IL15RA
CASP1
Interferon alpha
NOD2
FLT3LG
CXCL9
IL1B
TBX21
IgG
IL2RG
IL16
CXCR3
IL10
TP53
PF4
ERK
KIT
PDE4B
ERK1/2
PTN
CD3
TCR
STAT5a/b
PTPRZ1
Hsp27
P38 MAPK
CP: ERK/MAPK Signaling
FOXP3
TNFSF12
BMP6

(E)


LTB
IL1R1
Immunoglobulin
STAT4
CXCL11
NFkB (complex)
IL15RA
CASP1
Interferon alpha
NOD2
FLT3LG
CXCL9
IL1B
TBX21
IgG
IL2RG
IL16
CXCR3
IL10
TP53
PF4
ERK
KIT
PDE4B
ERK1/2
PTN
CD3
CP: SAPK/JNK Signaling
TCR
STAT5a/b
PTPRZ1
Hsp27
P38 MAPK
FOXP3
TNFSF12
BMP6

(F)

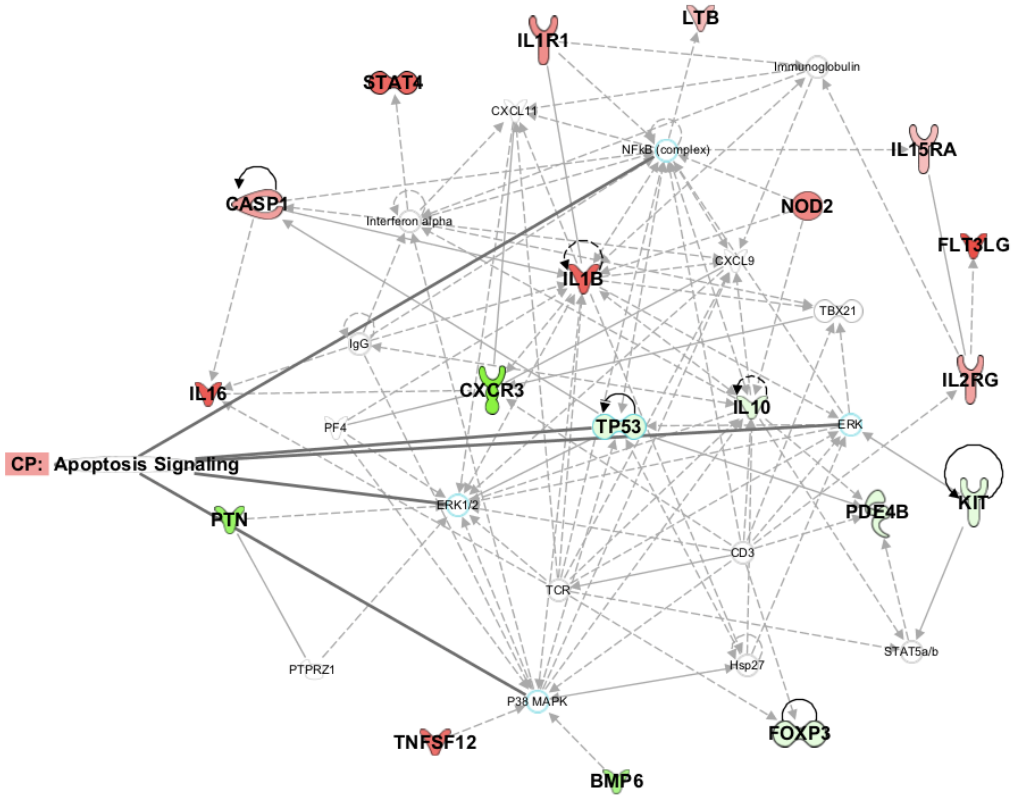


Figure 4.7. Connectivity of differentially expressed inflammatory diseases associated genes in colorectal cancer cells following HERV-H transduction. (A) Signalling network showing all involved genes. (B) Signalling network showing canonical pathway of molecular mechanisms of cancer. (C) Signalling network showing canonical pathway of colorectal cancer metastasis signalling (D) Signalling network showing canonical pathway of ERK/MAPK signalling. (E) Signalling network showing canonical pathway of SAPK/JNK signalling. (F) Signalling network showing canonical pathway of apoptosis signalling. Geometric figures in bold letters denote genes involvement in the study. Solid interconnecting lines shows the genes that are directly connected and the dotted lines signify indirect connection between the genes and cellular functions.

4.5.3 Western blot analysis of MAPK pathway

Based on the PCR array analysis, the protein levels of specific MAP kinases were examined. Western blot analysis using LS174T:HERV-H_GFP, LS174:GFP, HT29:HERV-H_GFP and HT29:GFP demonstrated that phosphor-p42/44 MAPK and phosphor-SAPK/JNK expressions were relatively higher in cells transfected with HERV-H (Figure 4.8).

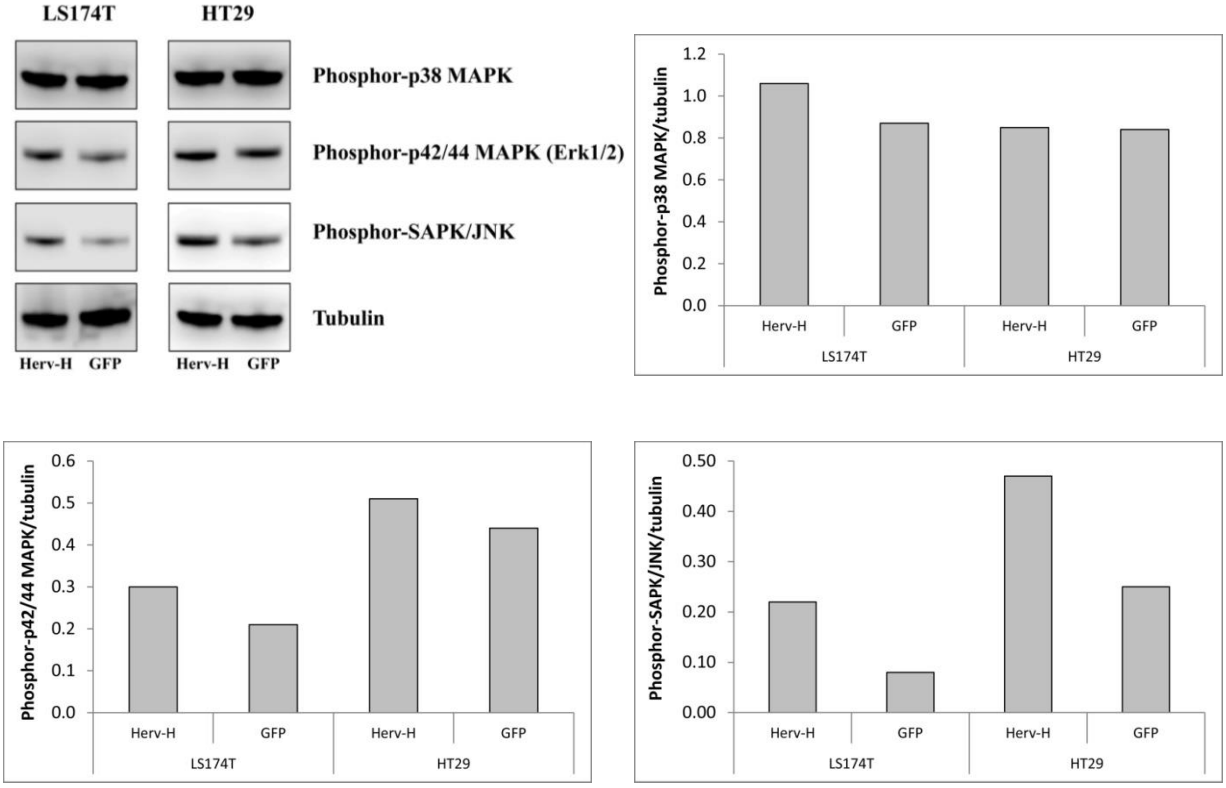


Figure 4.8. Western blot showing phosphor-MAP kinases in transfected colorectal cells. Left panel: LS174T cells transfected with GFP:HERV-H or GFP, right panel: HT29 cells transfected with GFP:HERV-H or GFP. Cells at 80% confluence were harvested and the protein concentrations of cell lysates were measured by Bradford assay. 40 μg of the cell lysate were loaded into each lane and subjected to immunoblotting. Immunoreactive bands for phosphor-p38 MAPK, phosphor-p42/44 MAPK, phosphor-SAPK/JNK and tubulin (housekeeping protein control) are shown. Band intensity was quantified using the Image Studio software (version 3.1.4) for C-DiGit Blot scanner (Li-Cor Biosciences, USA). The relative levels of phosphor-p38 MAPK, phosphor-p42/44 MAPK and phosphor-SAPK/JNK were normalised to their respective tubulin levels, and are shown.

4.6 Discussion

All cells possess the dynamic ability to coordinate their activities with environmental changes. In all multicellular organisms, such coordination depends on an elaborate network of communication that regulates the growth, differentiation, and metabolism of distinctive tissue types and organs. The basis of such communications is the engagement of signalling molecules, which can be used in either intercellular signalling or intracellular signalling (Krauss, 2008). The need to communicate and respond appropriately with the environment, neighbouring cells, as well as between intracellular compartments, is therefore a survival requirement (Sideman, 2005).

Interestingly, the very nature of embryonic development has engaged Hedgehog, Wnt (wingless related), transforming growth factor-β, receptor tyrosine kinase (RTK), Notch, Janus kinas (JAK) / signal transducer and activator of transcription (STAT) and nuclear hormone pathways for the tight spatial and temporal regulation of cell proliferation and differential (Pires-daSilva and Sommer, 2003). As such, cells are continuously exposed to and responding to a myriad of extracellular molecular signals and hence are tightly regulated in their lifetime. On this aspect, imbalances in signalling can lead to altered cell growth, differentiation and apoptosis (Dranoff, 2004).

In the preceding chapter, our study has demonstrated HERV-H endows colorectal cells with higher proliferative capacity, migratory ability and

reduced serum dependence. While the phenotypes were observable, the signalling pathways that HERV-H are involved in remains to be elucidated. Here, PCR array analysis, together with Western blot findings, revealed the implicated pathways when HERV-H-transfected colorectal cancer cells were compared with its own transfected control group of cells, GFP-transfected colorectal cancer cells.

Using GSEA and IPA application, PCR array analyses clearly indicated that there was a difference in the differential expression of genes and a probable mode of action underlying the disease and disorder development. While signalling pathway involving Wnt, PI3K-Akt and TGF-β were previously described in colorectal carcinogenesis (Fearon and Vogelstein, 1990, Takayama et al., 2006, Worthley et al., 2007, Markowitz and Bertagnolli, 2009, Ogino and Goel, 2008) , our study revealed the involvement of Wnt and MAPK (p42/44 or ERK1/2 and SAPK/JNK) when HERV-H is mediating the transforming process. Nevertheless, abrogation of p53 functions was consistently detected in both instances of colorectal cancer.

In the pathway analysis, IPA intriguingly associated the role of macrophages, fibroblasts and endothelial cells in rheumatoid arthritis (Table 4.5), a condition that could not be easily reconciled with carcinogenesis, in our study. A further examination revealed the up-regulation of various signalling pathways in the pathogenesis of rheumatoid arthritis. These pathways are the ERK/MAPK pathway, the NF-κB pathway, the Wnt pathway and the JAK-

STAT pathway. Activation of these pathways enhances the state of chronic inflammation with pro-inflammatory cytokines and growth factors, of which support the notion of cancer-related inflammation. Furthermore, macrophage migration inhibitory factor (MIF), which released from monocytes/macrophages, fibroblasts and endothelial cells (Babu et al., 2012, Calandra and Roger, 2003), is also found to play a role in colorectal carcinogenesis (He et al., 2009, Wilson et al., 2005, Conroy et al., 2010).

In the following sections, relevance of the uncovered signalling pathways to carcinogenesis will be discussed. These include the Wnt signalling pathway, the MAP kinase pathway, the p53 signalling pathway and the cytokines signalling pathways.

4.6.1 The Wnt signalling pathway

Wnt signalling plays a central role in embryonic development, differentiation, cell motility, cell proliferation, and adult tissue homeostasis. Wnt proteins or ligands are evolutionarily conserved, secreted cysteine-rich glycoproteins that transmit signals from outside a cell through cell surface receptors to the inside of the cell (Kikuchi et al., 2011). In humans, 19 members of the Wnt family are known (Papkoff et al., 1987).

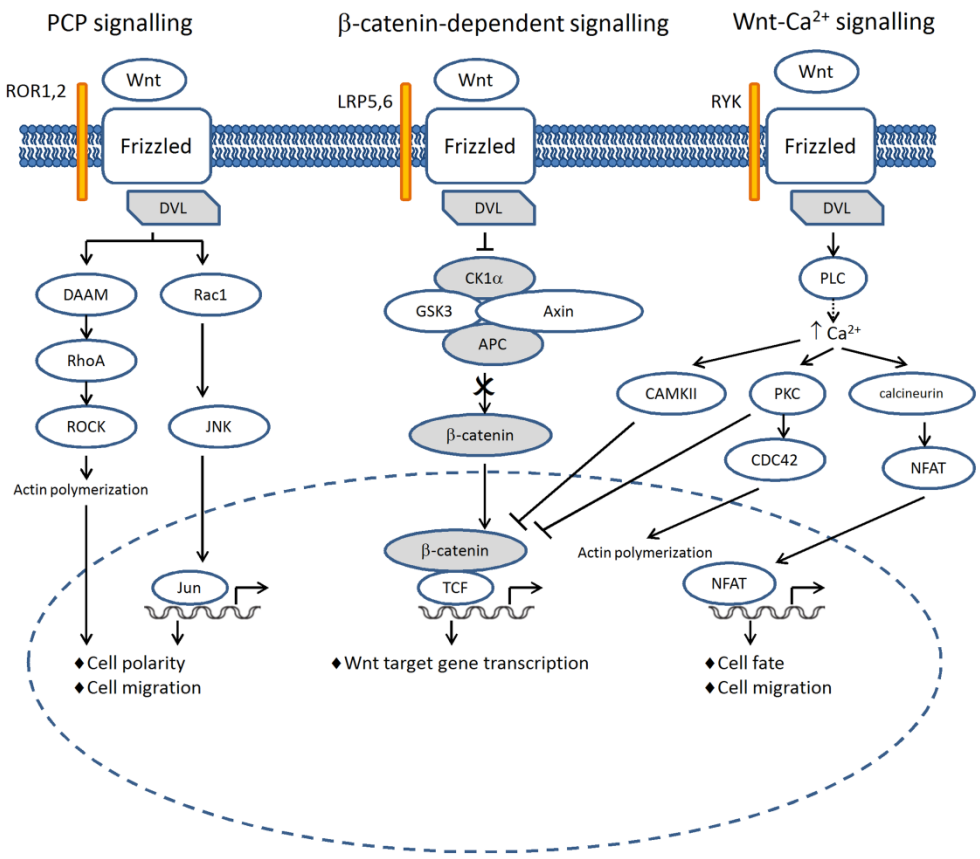


Figure 4.9. Wnt signalling pathways. Wnt regulates at least three major intracellular signalling pathways: the β-catenin-dependent pathway, the planar cell polarity (PCP) pathway and the Wnt-calcium pathway. PCP signalling triggers activation of the small GTPases RhoA and Rac1, which in turn activate Rho kinase (ROCK) and Jun-N-terminal kinase (JNK) respectively, resulting in actin polymerization and microtubule stabilization. This pathway is heavily involved in the regulation of cell polarity, cell motility and morphogenetic processes. In the β-catenin pathway, interaction of Wnt with Frizzled and LRP5 or LRP6 recruits and inactivates destruction complex that consists of glycogen synthase kinase 3 (GSK3), casein kinase 1α (CK1α), Axin and adenomatosis polyposis coli (APC). This allows β-catenin to accumulate and translocate to the nucleus, where it activates the transcription of target genes under the control of T cell factor (TCF). The Wnt-calcium pathway involves the activation of calcium- and calmodulin-dependent kinase (CAMKII), protein kinase C (PKC) and calcineurin. Calcineurin activates nuclear factor of activated T cells (NFAT), which modulates the gene transcription of cell fate and cell migration. CDC42: small GTPase; DAAM: DVL-associated activator of morphogenesis; DVL: Dishevelled; LRP: low-density lipoprotein receptor-related protein; PLC: phospholipase C; ROR: receptor tyrosine kinase-like orphan receptor; RYK: receptor tyrosine kinase. (Adapted from (Niehrs, 2012))

Based on conventional grouping, Wnt signalling pathways have been classified into canonical (β-catenin/CTNNB1-dependent) or non canonical (β-catenin/CTNNB1-independent) signalling pathways (Niehrs, 2012). Non canonical signalling pathways are further grouped as planar cell polarity (PCP) pathway and Wnt/Calcium pathway (Figure 4.9). Briefly, canonical Wnt signalling, mediated by Wnt ligands, regulates the destruction complex (comprising GSK3, CKIα, Axin and APC) to stabilise and translocate β-catenin to the nucleus where it regulates gene expression. In contrast to canonical Wnt signalling, non-canonical Wnt signalling, also mediated by Wnt ligands, regulates gene expression via the c-JUN-N-terminal kinase (JNK) or the calcium-dependent kinases, CAMKII and PKC.

It has been reported that Wnt1, Wnt3, and Wnt8 activate the canonical Wnt signalling, whereas Wnt5a and Wnt11 are thought to act mainly through the non-canonical Wnt pathway (Cohen et al., 2008). In addition, ten members of the frizzled family of G‑protein-coupled receptors, as well as the receptor tyrosine kinases (RTK) ROR1 and ROR2 and the RTK-like protein RYK, are known to be involved (Angers and Moon, 2009, Wang et al., 2006, Kohn and Moon, 2005).

The longstanding implication of the Wnt signalling pathway in carcinogenesis stems from the discovery of mouse proto-oncogene, *int1*, which is now known as *Wnt1* (Nusse and Varmus, 1982). Chronic activation of the Wnt signalling pathway has been associated with many human

malignancies including colorectal carcinomas and hepatocellular carcinomas (Logan and Nusse, 2004, Burgess et al., 2011).

Interestingly, recent studies have implicated tumour viruses are capable of modulating Wnt signalling and directing the dysregulated activation of Wnt pathway towards human cancer development. β-catenin, the central modulator of the canonical Wnt pathway, was found to be elevated in Epstein Barr virus-associated nasopharyngeal carcinoma (Hayward et al., 2006); nuclear accumulation of β-catenin in human papillomavirus infected cervical and oropharyneal cancer cells can also activate the Wnt pathway via its E6/E7 viral oncoproteins expression (Rampias et al., 2010).

4.6.1.1 Possible involvement of Wnt in HERV-H mediated transformation process

PCR array analyses revealed unique differential gene expression profiles that involved the signalling pathways and inflammation-association genes. Of interest, relatively higher expression of Wnt11, NFATC4 and DKK1 was observed in LS174T:HERV-H_GFP cells when compared to LS174T:GFP cells.

4.6.1.1.1 Wnt 11

Wnt11 is well established for its ability to trigger non canonical Wnt signalling (Pandur et al., 2002, Cha et al., 2008, Zhang et al., 2012). Specifically, Wnt11 signalling has been shown to promote intestinal cell proliferation, transformation and migration (Ouko et al., 2004) and that Wnt11 is upregulated in primary colorectal cancer (Nishioka et al., 2013). A recent study also revealed a higher Wnt11 mRNA expression in primary colorectal cancer tissues when compared to adjacent non-tumour tissues. In the same study, Wnt11 transfectants demonstrated increased phosphorylation of JNK and c-jun as well as increased proliferation and migration/invasion activities (Nishioka et al., 2013). The result of our study is consistent with this finding, suggesting that Wnt11/JNK signalling pathway is, at least partly, involved in the tumorigenesis process.

4.6.1.1.2 NFATC4

Nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 4 (NFATC4) is a member of the NFAT family consisting of NFATC1, NFATC2, NFATC3, NFATC4 and NFATC5, of all are activated via the control of calcineurin, a calcium dependent phosphatise. NFAT signalling, which is controlled by both negative and positive feedback (Crabtree and Olson, 2002), regulates the differentiation and development of the cardiovascular system and the musculoskeletal system as well as keratinocytes and adipocytes. Furthermore, NFAT signalling also plays a role in cell adaptation which aids cells like pancreatic cells, epidermal cells, and cardiac,

skeletal and smooth muscle cells to respond to the environmental changes (Horsley and Pavlath, 2002). NFAT has been closely linked with Wnt pathway (Wang et al., 2013, Niehrs, 2012, Sugimura et al., 2012, Sugimura and Li, 2010, Gregory et al., 2010). To date, numerous studies have implicated NFAT and Wnt signalling in colon cancers (Werneck et al., 2012, Slattery et al., 2011, Duque et al., 2005, Hong et al., 2007, Jauliac et al., 2002), including a few investigations that identifies NFATC2 as the critical player in colorectal carcinogenesis (Daniel et al., 2013, Gerlach et al., 2012). The increased expression of NFATC4 in LS174T:HERV-H_GFP cells compared to LS174T:GFP cells is consistent with these previous findings.

4.6.1.1.3 DKK1

While the introduction of HERV-H may perturb the signalling pathways in an upregulated manner, it is important to realise that a living cell is after all apt at homeostatic regulation, for instance, be programmed with negative feedback mechanisms. Thus, despite of the upregulation of genes involving proliferative activity after oncogenic Herv-H transduction, some genes are down-regulated.

As DKK1 is a potent Wnt signalling inhibitor (Glinka et al., 1998), by binding to LRP with high affinity and itself being as a target gene for signalling pathway (Bafico et al., 2001, Mao et al., 2001, Semenov et al., 2001) , DKK1 plays a role in modulating the Wnt pathway by establishing a

negative-feedback loop (Niida et al., 2004). Thus, a high DKK1 expression in our study indicates that the aberrantly upregulated Wnt signalling has resulted a corresponded increase in DKK1 expression.

4.6.2 The MAP Kinase pathway

The mitogen-activated protein kinase (MAPK) signalling pathway mediates a wide range of cellular activities which include cell proliferation, differentiation, cell survival, neuronal function and the immune response (Krishna and Narang, 2008). This large family of serine/threonine kinases, which can translocate into the nucleus, phosphorylates proteins like transcription factors, co-activators and co-repressors, is capable of regulating cellular transcriptional activities.

MAPK pathways are made up of a three-tier kinase module in which a MAPK is activated when phosphorylated by a mitogen-activated protein kinase kinase (MAPKK), which in turn is activated upon phosphorylation by a MAPKKK (Dhillon et al., 2007). The MAPK pathway is regarded as one of the most important pathways for cell proliferation since it entails major cell-proliferation signalling pathways from the cell surface to the nucleus (Fang and Richardson, 2005). In mammalian cells, the MAP kinases are classified into three main families, which include the ERKs (extracellular-signal-regulated kinases, ERK 1/2; also known as p42/44 MAP kinase), the JNKs (Jun amino-terminal kinases)/SAPKs (stress-activated protein kinases), and

the p38 MAP kinases (Morrison, 2012, Schaeffer and Weber, 1999). Generally, ERK 1 and ERK 2 are activated strongly by growth factors and are key transducers of proliferation signals. On the other hand, JNKs and p38 MAPK/SAPKs are poorly activated by growth factors but are strongly activated by cellular stress inducers like inflammatory cytokines, DNA damaging agents, oxidative stress, UV irradiation and growth factor deprivation (Figure 4.10) (Krishna and Narang, 2008).

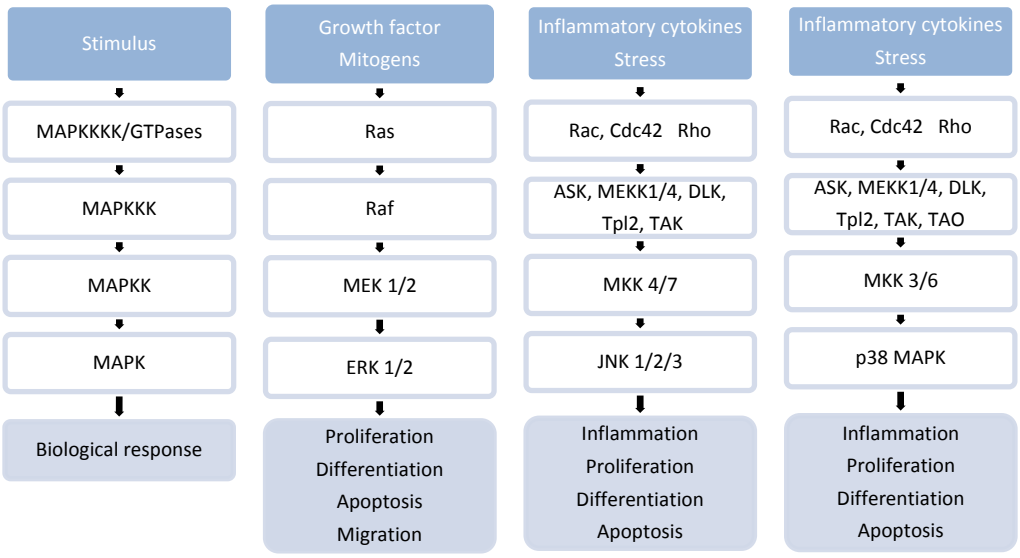


Figure 4.10. Schematic overview of MAPK pathways. MAPK signalling pathways mediate intracellular signalling initiated by extracellular or intracellular stimuli. MAPKKKs, which are activated by MAPKKKKs or GTPases, phosphorylate and activate MAPKKs, which in turn phosphorylate and activate MAPKs. Phosphorylated MAPKs then phosphorylate various substrate proteins like transcription factors and cofactors. This process regulates a variety of cellular activities which include cell proliferation, differentiation, migration, inflammatory responses and death. In mammalian cells, the MAPK family comprises ERKs, JNKs and p38.

It is widely accepted that MAPK signalling pathways are closely associated with the pathogenesis and progression of human colorectal cancer

(Fang and Richardson, 2005, Slattery et al., 2012, Lascorz et al., 2010). In contrast, blockade of MAPK pathway in human colorectal cancer cells can exert an antiproliferative effect on tumour cells (Sebolt-Leopold et al., 1999, Hoshino et al., 2001, Wang et al., 2004, Yeh et al., 2007, Sakamoto et al., 2013).

Studies showing viral activation of MAPK signalling pathways are accumulating. HBx, a 154-amino acid viral protein of 17 kDa encoded by hepatitis B virus, is a multifunctional regulator that is capable of modulating the cellular signal transduction pathways such as the Ras/Raf/MAPK pathway. Together with its role in upregulating proto-oncogenes like *c-jun*, *c-fos* and *c-myc*, HBx plays a significant role in the development of hepatocarcinoma (Motavaf et al., 2013, Ng and Lee, 2011, Matsuda and Ichida, 2009). Similarly, the Epstein-Barr virus latent membrane protein, LMP1, is found to mediate the activation of Ras/MAPK-dependent pathway (Roberts and Cooper, 1998) and the JNK/AP-1 cascade (Kieser et al., 1997, Hatzivassiliou et al., 1998), demonstrating its transforming potential in cells.

4.6.2.1 Possible involvement of MAPK (p42/44 or ERK1/2 and SAPK/JNK) in HERV-H mediated transformation process

Relatively lower expression of PIX2, MEF2C and MAP4K1 was found in LS174T:HERV-H_GFP cells when compared to LS174T:GFP. Our western blot finding also demonstrated a higher expression of phosphor-p42/44 and

phosphor-SAPK/JNK in LS174T:HERV-H_GFP cells as compared to LS174T:GFP cells.

4.6.2.1.1 PIX2

MAPK signalling has been closely linked to PIX2 in many studies (Basu and Roy, 2013, Sohn et al., 2009, Vadlamudi et al., 2005, Kioussi et al., 2002). Studies have also shown that PIX2 is implicated in breast cancer (Maier et al., 2007), ovarian cancer (Fung et al., 2012), prostate cancer (Hampton, 2006) and colorectal cancer, where overexpression of PIX2 is intriguingly associated with improved survival in patients as a result of reduced proliferative activity and invasiveness (Hirose et al., 2011). The reduced PIX2 expression in LS174T:HERV-H_GFP cells corroboratres with the findings demonstrated by Hirose *et al* where LS174T:HERV-H_GFP cells are indeed possessing a higher proliferative activity, albeit weak invasiveness.

4.6.2.1.2 MEF2C

Studies have shown that myocyte enhancer factor 2C (MEF2C) is closely associated with MAPK signalling pathway (Han et al., 1997, Wu et al., 2010, Khiem et al., 2008, Black and Olson, 1998). In addition, the oncogenic potential of MEF2C has been demonstrated not only in hepatocellular carcinoma (Bai et al., 2008), ovarian carcinoma (Kim et al., 2010), acute myelogenous leukaemia (Schwieger et al., 2009), T cell acute lymphoblastic leukaemia (Homminga et al., 2011, Nagel et al., 2008), and chronic myeloid

leukaemia (Agatheeswaran et al., 2013), but also in colorectal cancers (Nagaraj and Reverter, 2011) and gastric cancers (Ohta et al., 2009). Hence, a lower expression of MEF2C implies that the MAPK signalling pathway was modulated.

4.6.2.1.3 MAP4K1

The reduced expression of mitogen-activated protein kinase kinase kinase kinase 1 (MAP4K1) in our study is consistent with the finding that human colorectal adenocarcinomas, when compared to adult normal colon, has lost MAP4K1 in a cDNA microarray analysis (Kaiser et al., 2007). MAP4K1, also known as hematopoietic progenitor kinase 1 (HPK1), is a kinase upstream of JNK (MAP1K) (Ma et al., 2001). As such, MAP4K1 activates the JNK signalling pathway in a MAP4K1-TAK1-MKK4-JNK manner (Figure 4.6F) (Zhou et al., 1999). Specifically, MAP4K1 does not to implicate other MAPK signalling pathways like ERK and p38 signalling (Kiefer et al., 1996, Hu et al., 1996). While the signalling pathway of MAP4K1 has been mapped, its biological functions remain to be elucidated (Yang et al., 2006).

4.6.3 Possible co-involvement of Wnt and MAPK signalling pathways in HERV-H mediated transformation process

4.6.3.1 EGR1

Many previous studies have reported early growth response 1(EGR1) is regulated by the MAPK pathway (Grotegut et al., 2006, Tarcic et al., 2012, Xu et al., 2008, Lo et al., 2001, Saegusa et al., 2008). Hence, in our study, a high level of expressed EGR1 together with a high level of phosphor-p42/44 and phosphor-SAPK/JNK (Western blot) in LS174T:HERV-H_GFP implies the involvement of a MAPK signalling pathway (Figure 4.6E & F).

Interestingly, EGR1 gene is a transcription factor that acts as both a tumour suppressor and a tumour promoter. As a tumour suppressor, EGR1 binds to p53 for p53-mediated functions (Krones-Herzig et al., 2005). This mechanistic model, however, does not reconcile with our study.

On the other hand, the implication of EGR1 in the control of cell growth, survival and transformation emerged to be favourable (Thiel and Cibelli, 2002, Huang et al., 1998a). Specifically, the proliferative activity of EGR-1 was demonstrated in various human cancer models including prostate (Abdulkadir et al., 2001), skin (Riggs et al., 2000), and kidney(Scharnhorst et al., 2000). A variety of growth factors targeting at the EGR1 and mediating mitogenic signalling cascade were also identified in a microarray analysis (Svaren et al., 2000). More recently, it has been shown that *in vitro* EGR1-induced overexpression of TCF4 (transcription factor 4) resulted in the stabilization of nuclear β-catenin (in Wnt signalling pathway) and that may

play a role in the induction and maintenance of *trans*-differentiation in endometrial carcinoma cells (Saegusa et al., 2008).

Interestingly, one study has reported that the upregulation of EGR1 mRNA level was observed in the early-onset of colorectal cancers (Hong et al., 2007). As EGR1 is constitutively expressed in colon cancer cells, TRAIL-mediated apoptosis is inhibited. The inhibition effect is thought to be driven by the constitutive expression of c-FLIP (Mahalingam et al., 2010). Hence, the upregulation of EGR1 promotes the development of colon cancers.

Assimilating these findings, the Wnt/MAPK pathway emerged as the implicated signalling pathway during the HERV-H mediated transforming process.

4.6.4 The p53 signalling pathway

While p53 is described as the "guardian of the genome" (Lane, 1992) and the "cellular gatekeeper" (Levine, 1997), it is also known as a short-lived key tumour suppressor transcription factor in the cell (Levine et al., 2004, Vousden and Lane, 2007). The p53 tumour suppressor (or *TP53* gene) can be induced by a range of stresses including DNA damage, oncogene activation, or hypoxia (Horn and Vousden, 2007). Such p53 induction may lead to different biological outcomes such as apoptosis, cell-cycle arrest, senescence, or modulation of autophagy (Yee and Vousden, 2005, Riley et al., 2008,

Green and Kroemer, 2009, Jin and Levine, 2001). Among these, one of the most important functions will be innate tumour suppression (Meek, 2009).

The activation of the p53 protein in response to stresses is mediated and regulated by post-translational modifications which include phosphorylation, acetylation, methylation, ubiquitination or sumoylation. On the one hand, p53 can be activated by protein kinases, histone acetyltransferases, methylases, ubiquitin and sumo ligases; on the other hand, p53 can also be inactivated by phosphatases, histone deacetylases and ubiquitinases (Jin and Levine, 2001).

To exert its tumour suppressive properties, p53 mediates many biological effector functions (Figure 4.11). For instance, p53 is implicated in the cell cycle checkpoints, inducing G1 arrest mainly through the transactivation of p21 or inducing G2/M arrest mainly through the perturbation of the cyclin B1/cdc2 complex (Brugarolas et al., 1995, Deng et al., 1995).

p53 is also implicated in cellular senescence, in which p53 transactivates p21 via the association of p14ARF and Mdm2, and p21 inhibits E2F (a potent inducer of cell proliferation) via the inhibition of cyclin-dependent kinases upstream of the RB tumour suppressor.

In addition, p53 is implicated in autophagy, activating the damage-regulated autophagy modulator (*DRAM*) gene and inducing autophagy in a DRAM-dependent manner (Crighton et al., 2006). Similarly, p53 is implicated in apoptosis, either transactivating many genes involved in apoptosis, which include *Bax*, *PIG3*, *Killer/DR5*, CD95 (Fas), *p53AIP1*, *Perp*, and BH3-only proteins Noxa and PUMA (p53-up-regulated modulator of apoptosis) (Riley et al., 2008), and/or inducing mitochondrial outer membrane permeabilization (MOMP) as well as interacting with *Bcl2*, *Bcl-X_L*, and *Bak* at the mitochondria (Green and Kroemer, 2009).

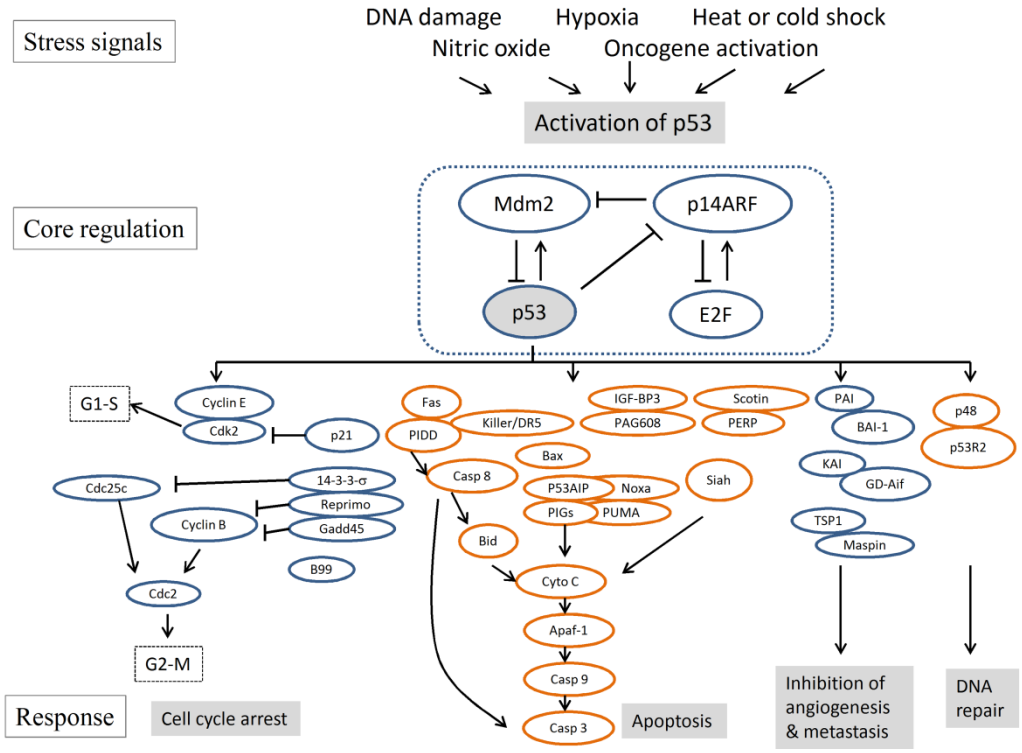


Figure 4.11. The p53 functional circuit. The diversity of stress signals that activate the p53 protein contributes to the central role of p53 as a tumour suppressor. The downstream targets of p53 mediating the different biological outcomes like cell cycle arrest, apoptosis, senescence and DNA repair are shown. (Adapted from (Harris and Levine, 2005)

More than 38% of almost all human cancers contain somatic *TP53* mutations (Olivier et al., 2010). Consequently, p53 function is often inactivated or suppressed in human cancers (Green and Kroemer, 2009). The inactivation of the p53 pathway by mutation of *TP53* is one of the key genetic molecular aetiology in colorectal cancer (Markowitz and Bertagnolli, 2009). In many cases of inactivated *TP53*, the transition of large adenomas into invasive carcinomas is often encountered (Baker et al., 1990).

Of particular interest, various DNA viruses, such as simian virus 40 (SV40), human papillomavirus (HPV), adenoviruses or Epstein-Barr virus (EBV), encode proteins that target the p53 protein for inactivation. For instance, SV40 encodes a large T-antigen (Tag) that binds to p53 and inactivates its effector functions via its blockade of p53 binding to the ribosomal gene cluster (RGC) site (Bargonetti et al., 1992).

On the other hand, the E6 proteins of HPV-16 and -18 bind to the p53 protein and inactivate it by accelerating its proteolytic degradation via poly-ubiquitination (Scheffner et al., 1990, Scheffner et al., 1993, Munger and Howley, 2002).

Likewise, the E1B55K and E4orf6 proteins of adenovirus are capable of forming a complex with the p53 protein, and thereby recruiting a Cullin-

containing complex to direct the ubiquitin-mediated proteolysis of p53 (Sarnow et al., 1982, Querido et al., 2001).

Similarly, the BZLF1 protein from EBV is capable of reconstituting a multiprotein ECS (*E*longin B/*C*-Cul2/5-*S*OCS-box protein) complex with ubiquitin ligase activity, and this complex targets p53 for ubiquitination and degradation (Sato et al., 2009). Taken together, viral proteins are capable of undermining p53 functions and thus contribute to the progression of cancers (Table 4.6).

Table 4.6. Viruses that act on p53 inactivation (Collot-Teixeira et al., 2004)

Virus	Associated human cancer	p53 binding
Adenovirus	Not documented	E1B55K/E4orf6
SV40	Not documented	Large T antigen
JC and BK polyomaviruses	Brain tumours	Large T antigen
EBV	Burkitt's lymphoma	BZLF1
HPV (16/18)	Cervical carcinoma	E6

Each of the DNA tumour viruses encodes oncogene products that associate with p53 for inactivation.

4.6.4.1 Abrogation of p53 in HERV-H mediated transformation process

It is crucial to realise that it is the wildtype *TP53*, but not its mutant forms, is functioning as a tumour suppressor gene (Eliyahu et al., 1989, Finlay et al., 1989). Thus, the lowered *TP53* expression found in LS174T:HERV-H_GFP cells must be evaluated in consideration of *TP53* sequences integrity.

Consistent with ATCC colon cancer p53 hotspot mutation cell panel (ATCC® TCP-2020™) specifications, a previous study has shown that LS174T cell line possesses a wildtype *TP53* (Liu and Bodmer, 2006). Given these information, a lowered expression of wildtype *TP53* indicates a greater degree of tumorigenicity of LS174T:HERV-H_GFP colorectal cells (Figures 4.5 & 4.7F). This finding corroborates with a recent study that demonstrates *TP53* mRNA expression was lower in tumours when compared to paired non-neoplastic specimens (Calcagno et al., 2013).

A decreased expression of cell death-inducing p53 target 1 (CDIP) or C16orf5 was also found in LS174T:HERV-H_GFP cells. CDIP was previously demonstrated to be a novel p53 target that induced apoptosis via the intrinsic apoptotic pathway. As such, a reduced expression of CDIP indicates that there is a lowered inhibition of CDIP-mediated abrogation of p53-mediated apoptosis (Brown et al., 2007).

Comparatively low expression of v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog (KIT) was likewise demonstrated in LS174T:HERV-H_GFP cells (Figure 4.7A). This finding corroborates with a previous study that shows c-Kit was suppressed in human colorectal cancer cells (Gavert et al., 2013). In addition, a more aggressive cancer phenotype resulting from the loss of c-Kit has also been demonstrated in many cancers (Maffini et al., 2008, Tsutsui et al., 2006, Tonary et al., 2000, Zakut et al., 1993) , including melanoma, where c-Kit expression in a highly metastatic

cell line resulted in a reduction of tumorigenicity and metastatic capacity (Huang et al., 1998b).

4.6.5 Inflammation: cytokines signalling

Inflammation is a physiological host response to tissue damage following microbial pathogen infection, chemical irritation, or physical insult (Sgambato and Cittadini, 2010). An inflammatory response can be acute or chronic.

Acute inflammation is a short-term response to the triggering stimuli and is characterized by leukocytes (neutrophils) infiltrating at the site of infection or injury with roles to remove obnoxious stimuli and carry out tissue repair (Ryan and Majno, 1977, Medzhitov, 2008). On the other hand, chronic inflammation is a prolonged and dysregulated response characterised by active inflammation, domination of macrophages, T lymphocytes and plasma cells, angiogenesis, tissue destruction and fibrosis (Kumar et al., 2009). Whereas acute inflammation normally impedes the development of cancer, chronic inflammation promotes cancer development (Coussens and Werb, 2002, Philip et al., 2004).

In fact, it was in 1863 that Rudolf Virchow put forward the idea that the origin of cancer was at sites of chronic inflammation (Balkwill and Mantovani, 2001). On the same note, it was postulated that within such a rich microenvironment of biological mediators like cytokines, chemokines,

interleukins, inducible cyclooxygenase (COX-2), inducible nitric oxide synthase (iNOS/NOS2), and matrix metalloproteinases (MMPs) (Aggarwal et al., 2006), this provides a good platform for the development of cancer as it encompasses multiple signalling pathways that corroborate with the hallmarks of cancer (Hanahan and Weinberg, 2011). To this end, the inflammatory microenvironment serves to provide an inflammatory cytokine network that modulates survival, growth, mutation, proliferation, differentiation and movement of both tumour and stromal cells (Balkwill and Mantovani, 2001), ssupporting the notion of cancer-related inflammation as the seventh hallmark of cancer in the past (Colotta et al., 2009).

The release of cytokines is in response to a wide range of cellular stresses. Consequently, normal physiological host responses to these released cytokines aid in containing the associated biological damages while promoting recovery. When such a homeostatic mechanism fails, persistent production of cytokines ensues. This leads to the dysregulation of cellular signalling and erroneously promotes neoplastic programming of the affected cells (Johansson et al., 2008).

The association between inflammation and cancer has been mapped as the convergence of two pathways: the extrinsic pathway which is driven by inflammatory conditions and the intrinsic pathway which is driven by activated oncogenic events (Figure 4.12) (Mantovani et al., 2008). The role of

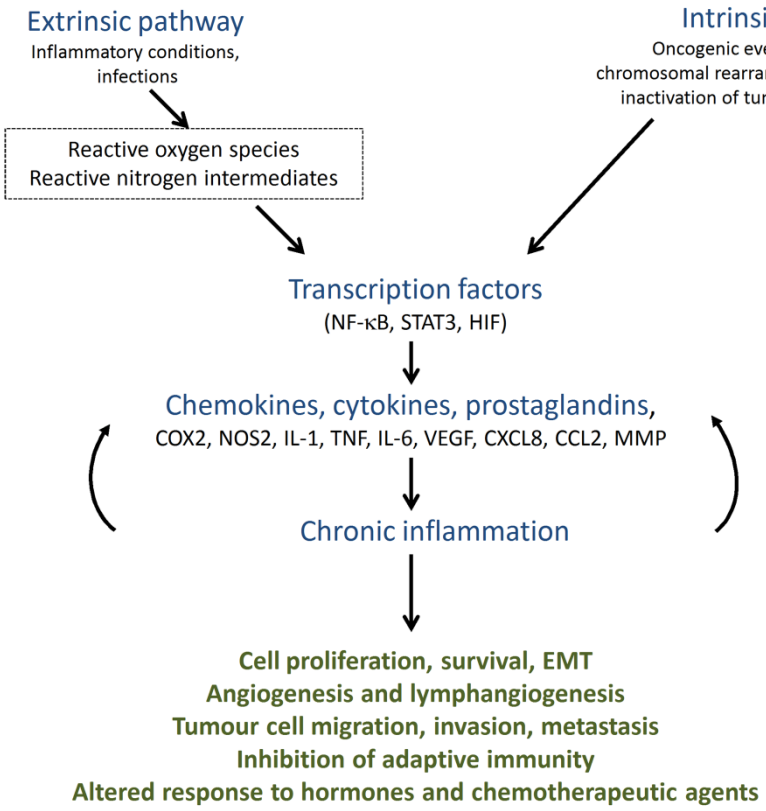


Figure 4.12. Pathways that link inflammation and cancer. A schematic view of the link between chronic inflammation and cancer via two pathways: the intrinsic pathway and the extrinsic pathway. The intrinsic pathway is driven by oncogenetic events like mutations that cause neoplasia. The extrinsic pathway is driven by inflammatory conditions and infections. Induced by inflammation, reactive oxygen species and nitrogen intermediates are capable of DNA damage. Activation of transcription factors, predominantly nuclear factor-κB (NF-κB), signal transducer and activator of transcription 3 (STAT3) and hypoxia-inducible factor (HIF), in tumour cells occurs after the two pathways converge. Resultant effects include the production of cytokines, chemokines and prostaglandins as well as the recruitment and activation of immune cells (chronic inflammation). Activated immune cells secrete more cytokines and provide an inflammatory microenvironment that has tumour-promoting effects. COX2: cyclooxygenase 2, NOS2: inducible nitric oxide synthase, IL-1: interleukin 1, TNF: tumour necrosis factor, IL-6: interleukin 6, VEGF: vascular endothelial growth factor, CXCL8: CXC-chemokine ligand 8, CCL2: CC-chemokine ligand 2, MMP: matrix metalloproteinase, EMT: epithelial–mesenchymal transition. (Adapted from (Mantovani et al., 2008, Aggarwal et al., 2006)).

cytokines, chemokines and interleukins involved in the tumorigenesis is interestingly complex but has been summarized in Table 4.7. In general, pro-inflammatory cytokines are pro-apoptotic and anti-inflammatory cytokines are anti-apoptotic (Schuerwegh et al., 2003). The role of immune cells in tumour-promoting inflammation is presented in Table 4.8.

Table 4.7. Role of cytokines and other molecular players in tumorigenesis
(Terzic et al., 2010, Lu et al., 2006, Aggarwal et al., 2006, Dranoff, 2004)

Agent	Mechanisms
TNFα	Promote survival, activation, recruitment, growth. AP-1, MAPK and NF-κB activation
IL-6	Promote survival, growth, STAT3, ERK, and Akt activation
IL-11	Promote survival, growth. STAT3, STAT1, ERK activation
IL-23	T-cell differentiation (Th17) and interference with Tregs, production of IL-17 and IL-22 by immune cells.
IL-15	Promotes proliferation, survival, prevents apoptosis
IL-1α, IL-1β	Survival, growth, cytokines, chemokines, tumour invasion, angiogenesis. NF-κB, MAPK activation
IL-22	Survival, mucosal integrity, chemokines. STAT3 activation
IL-17A,F	Survival, chemokines, T-cell regulation, monocytes, and neutrophil recruitment. MAPK, NF-κB activation
IL-8	Promote tumour growth, angiogenesis
IL-18	Metastasis
Epidermal growth factor	Survival, proliferation. MAPK, STAT3 activation

IL-10	Anti-inflammatory, Treg stimulation. STAT3, MAPK activation
macrophage-migration inhibitory factor	inhibits p53 tumour-suppressor functions
Chemokines	Promote tumor cell growth Facilitate invasion and metastasis by directing tumor cell migration and promoting basement membrane degradation
NF-κB	Mediate inflammation progress, promoting chronic inflammation. Promote the production of mutagenic reactive oxygen species. Protect transformed cells from apoptosis. Promote tumor invasion and metastasis. Feedback loop between proinflammatory cytokines
iNOS	Downstream of NF-κB and proinflammatory cytokines. Induce DNA damage and disrupt DNA damage response. Regulate angiogenesis and metastasis
COX-2	Produce inflammation mediator prostaglandins. Promote cell proliferation, antiapoptotic activity, angiogenesis, and metastasis
HIF-1α	Promote chronic inflammation. Induced by proinflammatory cytokines through NF-κB. Enhance the glycolytic activity of cancer cells. Contribute to angiogenesis, tumor invasion, and metastasis by transactivating VEGF
STAT3	Activated by proinflammatory cytokines. Promote proliferation, apoptosis resistance, and immune tolerance
Nrf2	Anti-inflammatory activity. Protect against DNA damage.
NFAT	Regulate proinflammatory cytokine expression. Required in cell transformation
CXCR4	Metastasis, invasion and growth, proliferation

CCR7	Metastasis
CCR10	Metastasis
CCR6	Cell invasion
MIP-3a	Cell invasion
SDF1	Proliferation
ERK	Proliferation
AKT	Proliferation

Table 4.8. Role of immune cells in tumorigenesis (Grivennikov et al., 2010, Terzic et al., 2010)

Cell types	Tumour-promoting
Macrophages, dendritic cells, myeloid-derived suppressor cells	Immunosuppression; production of cytokines (IL-6, IL-1,VEGF, IL-23,TNF), chemokines, proteases, growth factors and angiogenic factors
Mast cells	Production of cytokines
B cells	Production of cytokines and antibodies; activation of mast cells; immunosuppression
CD8+ T cells	Production of cytokines?
CD4+ Th2 cells	Education of macrophages; production of cytokines; B cell activation
CD4+ Th1 cells	Production of cytokines
CD4+ Th17 cells	Production of cytokines
CD4+ Treg cells	Immunosuppression; production of cytokines (IL-10, TGF-β) Suppress inflammation
Neutrophils	Production of cytokines, proteases, and ROS
NK cells	Production of cytokine production (IFNγ, IL-22, IL-17)

Chronic viral infection is gradually being recognised as a leading cause of cancer worldwide (Parkin, 2006, Zur Hausen, 2009, de Martel et al., 2012). With HPV, HBV, EBV, HTLV-1, HCV, KSHV implicated in the process of tumorigenesis (Mattock, 2012), chronic inflammation associated with chronic viral infection has provided a strong link to viral mediated carcinogenesis (McLaughlin-Drubin and Munger, 2008, Zhang et al., 2013b, Butel, 2000) . In such viral-mediated carcinogenesis, signalling pathways involving NF-κB, STAT3 and hypoxia-inducible factors (HIFs) are often implicated (Mantovani et al., 2008) (Figure 4.12).

4.6.5.1 Possible involvement of inflammation-associated cytokine signalling in HERV-H mediated transformation process

LS174T:HERV-H_GFP colorectal cells showed relatively higher expression in interleukin 1B (IL1B), interleukin 1 receptor, type I (IL1R1) and caspase 1 (CASP1) when compared to LS174T:GFP.

4.6.5.1.1 IL1B

IL1B is a highly active proinflammatory cytokine that has a pivotal role in autoinflammatory diseases, a diverse group of disorders characterised by abnormal activation of the innate immune system and chronic inflammation with high levels of acute-phase reactants (Ozen and Bilginer, 2013). Together with its angiogenic capability (Voronov et al., 2003) and ability to recruit immune cells (Rider et al., 2011), IL1B is found to play a role

in tumour metastasis (Dinarello, 2009). In certain cancers like melanoma (Li et al., 2009), acute myelogenous leukemia (Giavazzi et al., 1995) and multiple myeloma(Lust and Donovan, 1999), IL1B is expressed as part of their oncogenic nature. In line with previous studies, higher expression of IL1B was observed in LS174T:HERV-H_GFP cells when compared with LS174T:GFP cells.

4.6.5.1.2 IL1B, CASP1 and IL1R1

Caspase-1 is an enzyme that is involved in the processing of pro-IL1B to active, secreted IL1B. Thus, the consequence of caspase-1 activation is the secretion of IL1B. Furthermore, IL1B regulates its own production and processing (Denes et al., 2012). This suggests that the concordance rise in IL1B and CASP1 expression in LS174T:HERV-H_GFP might well correspond to the inflammatory state that the transfected cells was in, given that HERV-H is after all a viral peptide (Figure 4.6A & 4.7A).

On the same note, the surface expression of IL1R1 is promoted by the inflammatory mediators like prostaglandin E2, epidermal growth factor, IL-2 and IL-4 (Dinarello, 2009). As LS174T:HERV-H_GFP cells might well be in an inflammatory microenvironment, the existence of a myriad of cytokines may collectively enhance the expression of IL1R1.

4.6.5.1.3 IL-16

Compared to LS174T:GFP cells, interleukin 16 (IL-16) was found to be upregulated in LS174T:HERV-H_GFP cells. IL-16, originally identified as a lymphocyte chemoattractant factor in 1982 (Center and Cruikshank, 1982), is currently regarded as a proinflammatory cytokine (Cruikshank et al., 2000), although it does not induce detectable apoptotic cell death (Krautwald, 1998). Previous study has shown that the surface expression of IL-2Rα and IL-2Rβ was increased following IL-16 stimulation. However, there was no increase in IL-2Rγ expression (Parada et al., 1998). In contrast to this finding, our data showed increased expression of IL-2Rγ (Figure 4.5), suggesting that the effect of IL-16 might be cell specific (CD4+ T cells versus colorectal cells). Nevertheless, IL-16 is an activator of the SAPK signalling pathway (Figure 4.7E) (Krautwald, 1998). While the correlation between IL-16 and tumour growth is limited by available literature, studies have shown that IL-16 was lower in treated non-Hodgkin lymphomas patients (Passam et al., 2008), and higher in astrocytic brain tumours (Liebrich et al., 2007) and multiple myeloma (Liebrich et al., 2007). In addition, increased IL-16 mRNA expression was also found in cutaneous T cell lymphoma (Blaschke et al., 1999).

4.6.6 Concluding remarks

Taken together, HERV-H transfected LS174T cells may represent a feasible *in vitro* model to study the signalling pathways involved in endogenous viral-mediated carcinogenesis. Our data has shown that pathways

involving p42/44 MAPK, SAPK/JNK as well as p53 may be implicated in HERV-H-mediated oncogenic mechanism. These findings have important implications for the development of rationale therapeutics, as it raises the possibility to target signalling pathways that are highly active specifically in colorectal cancer cells.

General Discussion and Conclusion

Chapter 5

CHAPTER 5

GENERAL DISCUSSION AND CONCLUSION

5.1 Viral-mediated carcinogenesis

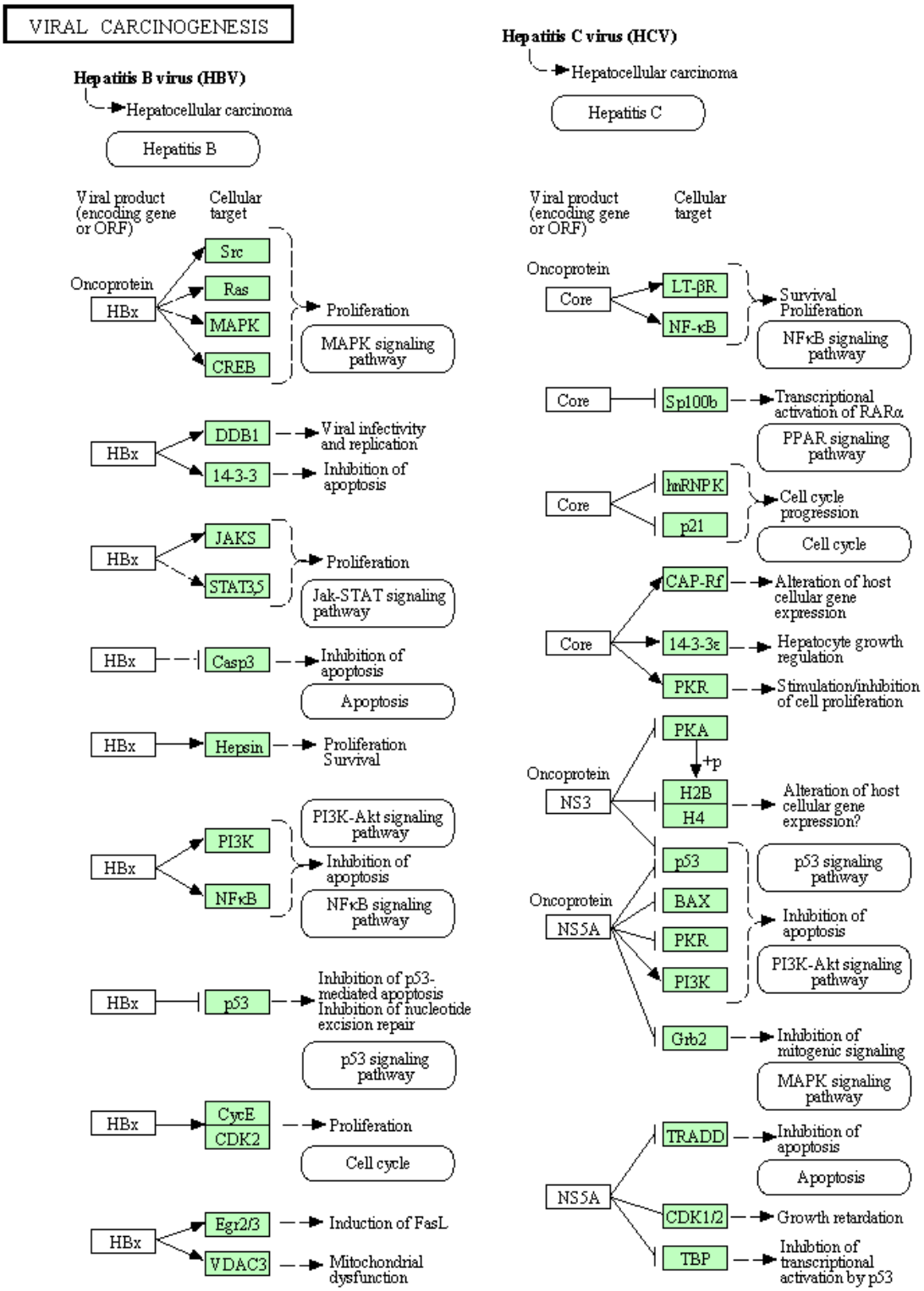
The concept of viral mediated carcinogenesis has been well established (Haverkos, 2004, Butel, 2000, Carrillo-Infante et al., 2007, McLaughlin-Drubin and Munger, 2008, Zhang et al., 2013b). A comprehensive account of the diverse aspects of viral carcinogenesis is intricately arduous. Thus, only six oncogenic viruses that are classified by IARC as "carcinogenic to humans" (Group I) will be briefly discussed. As previously mentioned in Chapter 1, these six oncogenic viruses include four DNA viruses – EBV, HBV, HPV and KSHV and two RNA viruses – HCV and HTLV-1.

All tumour viruses are capable of modulating signalling pathways via their oncogene products to regulate cellular proliferation, migration, apoptosis, differentiation and recognition by the immune system. An overview of the viral oncoproteins together with their associated cellular targets and deregulated signalling pathways is presented in Table 5.1. In addition, a KEGG pathway map depicting the viral carcinogenesis (Kanehisa, 2013, Kanehisa et al., 2012) is illustrated in Figure 5.1. All these findings illuminate the infectious aetiology of certain human cancers.

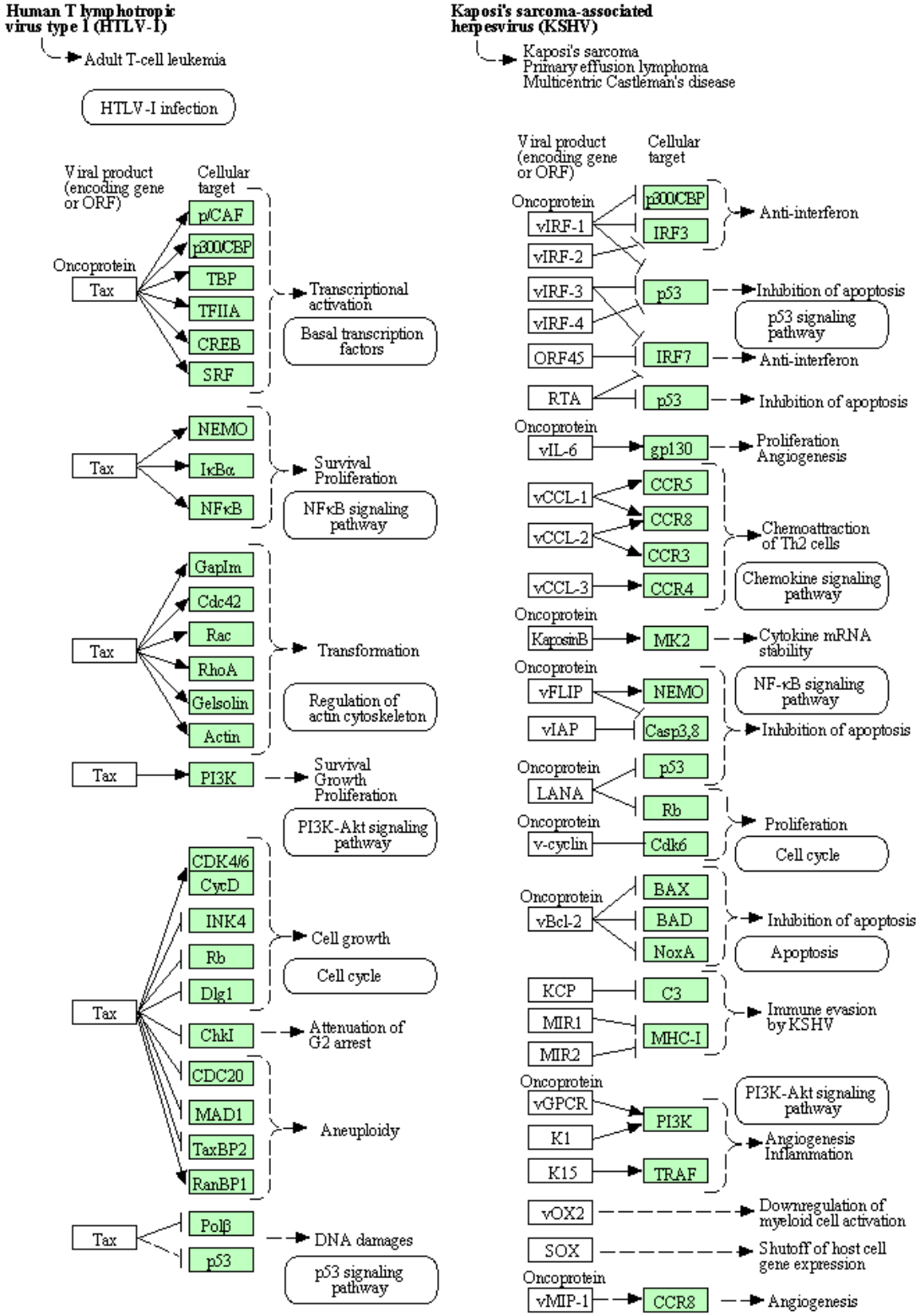
Table 5.1. Oncogenic virus and associated cell-signalling molecules (Adapted from (Saha et al., 2010, Butel, 2000)).

Oncogenic virus	Viral genome	Integrated viral genome in tumours	Viral oncoproteins	Cellular targets	Deregulated signalling pathways
HPV16, 18	DNA	Yes, usually	E6	p53, p73, E6AP, CBP/p300, c-Myc	p53, cell-cycle
			E7	pRb, pRb pocket proteins, $p21^{CIP1}$, $p27^{KIP1}$, IRF-1, Cyclin A and E	pRb, cell-cycle, ub-proteasome
			EBNA2	RBP-Jκ, PU.1, AUF1, DDX20, SMN	Notch, cellular transcription, metastasis
			EBNA3C	p53, Mdm2, pRb, p300, RBP-Jκ, Chk2, Nm23-H1, c-Myc, HDAC1, SUMO-1, SUMO-3, SCF^{Skp2}-complex, DDX20, SMN, CtBP, Cyclin A, E and D1	Cell-cycle, Notch, ub-proteasome, metastasis, chromatin remodeling, cellular transcription, apoptosis, inflammation
EBV	DNA	-	LMP1	TRAFs 1, 2, 3 and 5, TRADD, RAS, JAK	NFκB, cell-cycle, cellular transcription, apoptosis, inflammation, autophagy, metastasis, MAPK, PI3K/Akt, JAK/STAT, TNF
			LMP2	TNFR associated factors, RAS, JAK	Apoptosis, metastasis, MAPK, PI3K/Akt, JAK/STAT, TNF, BCR signalling
KSHV	DNA	-	LANA	p53, pRb, c-Myc, GSK3β, MAPK2, FADD, core histones, Transcriptional activators- Brd2, Brd4, Sp1, AP-1, and CBP and transcriptional inhibitors HP1, Dnmt3 and mSin3	cell-cycle, cellular transcription, apoptosis, Notch, Wnt/β-catenin, ub-proteasome, chromatin remodeling
			vFLIP	TRAF2	Apoptosis, NFκB, JNK/AP1
HBV	DNA	Yes, usually	HBx	NFκB, p53, c-jun, c-fos, PKC, c-Myc, SP1, HIF-1α	Cell-cycle, apoptosis, cellular transcription, NFκB Wnt/β-catenin, TGFβ, JAK/STAT, metastasis
HCV	RNA	No	NS3	p53, Arginine methyltransferase 1, PKA, H2B, H4	PKC, inflammation
			NS5A	p53, Bax, IFN-induced dsRNA activated protein kinase (PKR), growth factor receptor-binding protein 2 (Grb2), PI3K p85 subunit, TRADD, CDK1, TRAF2, TBP	Cell-cycle, apoptosis, Ras-Erk MAPK pathway, PI3K, NFκB
HTLV-1	RNA	Yes, as provirus	Tax	Cyclic AMP, p300/CBP, MAD-1, MAD-2, cyclin D1, Chk1 and 2	Cell-cycle, apoptosis, cellular transcription, NFκB, PI3K/AKT, chromatin remodelling

(A)


VIRAL CARCINOGENESIS
Hepatitis B virus (HBV)
Hepatocellular carcinoma
Hepatitis B
Viral product (encoding gene or ORF)
Cellular target
Oncoprotein
HBx
Src
Ras
MAPK
CREB
Proliferation
MAPK signaling pathway
DDB1
Viral infectivity and replication
14-3-3
Inhibition of apoptosis
JAKS
STAT3,5
Proliferation
Jak-STAT signaling pathway
Casp3
Inhibition of apoptosis
Apoptosis
Hepsin
Proliferation Survival
PI3K
NFκB
PI3K-Akt signaling pathway
Inhibition of apoptosis
NFκB signaling pathway
p53
Inhibition of p53-mediated apoptosis
Inhibition of nucleotide excision repair
p53 signaling pathway
CycE
CDK2
Proliferation
Cell cycle
Egr2/3
Induction of FasL
VDAC3
Mitochondrial dysfunction
Hepatitis C virus (HCV)
Hepatocellular carcinoma
Hepatitis C
Core
LT-βR
NF-κB
Survival Proliferation
NFκB signaling pathway
Sp100b
Transcriptional activation of RARα
PPAR signaling pathway
hnRNPK
p21
Cell cycle progression
Cell cycle
CAP-Rf
Alteration of host cellular gene expression
14-3-3ε
Hepatocyte growth regulation
PKR
Stimulation/inhibition of cell proliferation
PKA
+p
NS3
H2B
H4
Alteration of host cellular gene expression?
p53
p53 signaling pathway
NS5A
BAX
PKR
PI3K
Inhibition of apoptosis
PI3K-Akt signaling pathway
Grb2
Inhibition of mitogenic signaling
MAPK signaling pathway
TRADD
Inhibition of apoptosis
Apoptosis
CDK1/2
Growth retardation
TBP
Inhibition of transcriptional activation by p53

(B)


Human T lymphotropic virus type I (HTLV-I)
Adult T-cell leukemia
HTLV-I infection
Viral product (encoding gene or ORF)
Cellular target
Oncoprotein
Tax
p/CAF
p300/CBP
TBP
TFIIA
CREB
SRF
Transcriptional activation
Basal transcription factors
NEMO
IκBα
NFκB
Survival Proliferation
NFκB signaling pathway
GapIm
Cdc42
Rac
RhoA
Gelsolin
Actin
Transformation
Regulation of actin cytoskeleton
PI3K
Survival Growth Proliferation
PI3K-Akt signaling pathway
CDK4/6
CycD
INK4
Rb
Dlg1
Cell growth
Cell cycle
ChkI
Attenuation of G2 arrest
CDC20
MAD1
TaxBP2
RanBP1
Aneuploidy
Polβ
p53
DNA damages
p53 signaling pathway
Kaposi's sarcoma-associated herpesvirus (KSHV)
Kaposi's sarcoma
Primary effusion lymphoma
Multicentric Castleman's disease
Viral product (encoding gene or ORF)
Cellular target
Oncoprotein
vIRF-1
vIRF-2
vIRF-3
vIRF-4
ORF45
RTA
p300/CBP
IRF3
IRF7
p53
Anti-interferon
Inhibition of apoptosis
p53 signaling pathway
vIL-6
gp130
Proliferation Angiogenesis
vCCL-1
vCCL-2
vCCL-3
CCR5
CCR8
CCR3
CCR4
Chemoattraction of Th2 cells
Chemokine signaling pathway
KaposinB
MK2
Cytokine mRNA stability
vFLIP
vIAP
NEMO
Casp3,8
NF-κB signaling pathway
Inhibition of apoptosis
LANA
v-cyclin
p53
Rb
Cdk6
Proliferation
Cell cycle
vBcl-2
BAX
BAD
NoxA
Apoptosis
KCP
MIR1
MIR2
C3
MHC-I
Immune evasion by KSHV
vGPCR
K1
K15
PI3K
TRAF
PI3K-Akt signaling pathway
Angiogenesis Inflammation
vOX2
Downregulation of myeloid cell activation
SOX
Shutoff of host cell gene expression
vMIP-1
CCR8
Angiogenesis

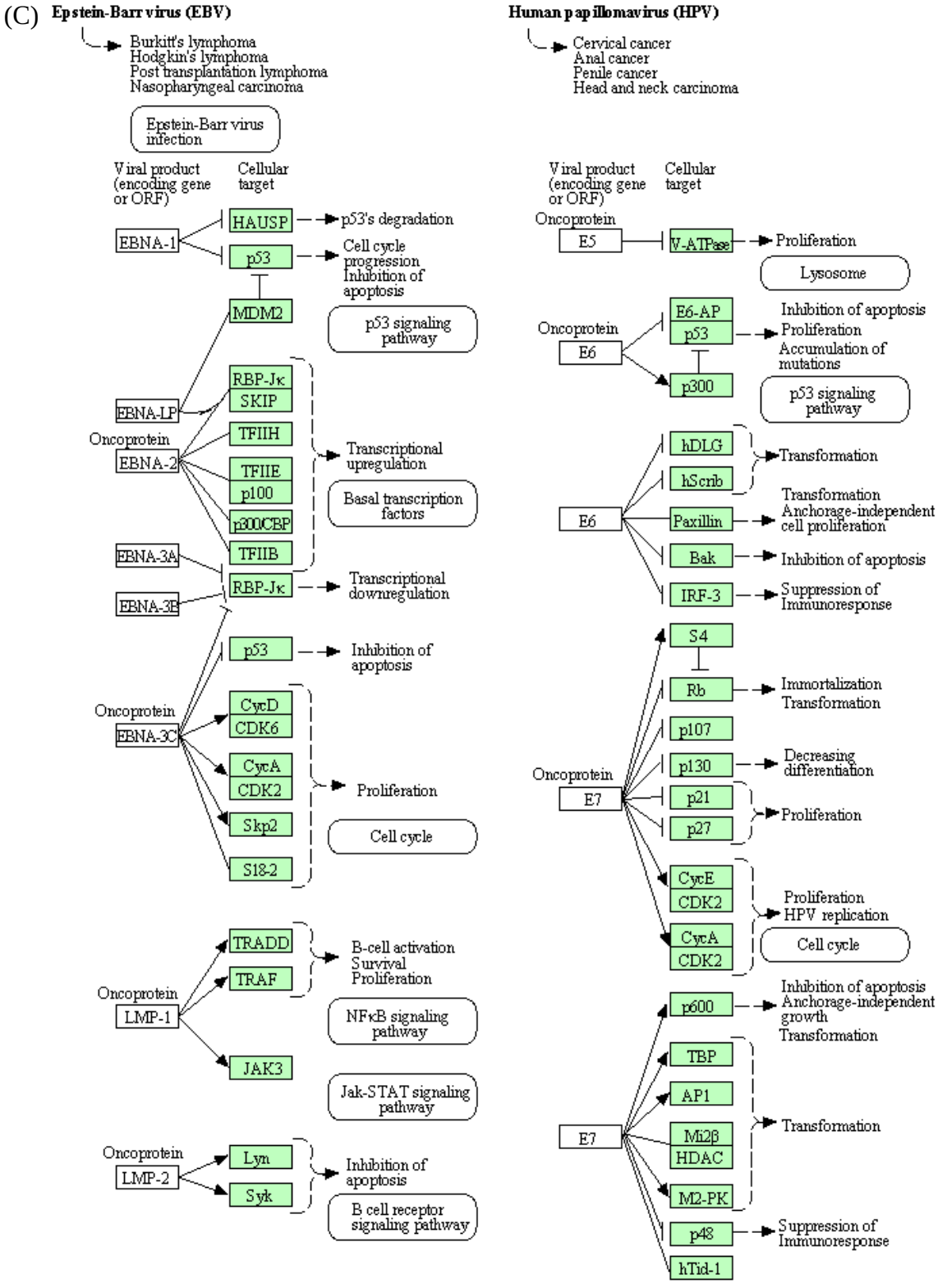


Figure 5.1. KEGG pathway map showing the deregulated pathways in viral carcinogenesis. Viral carcinogenesis involving (A) hepatitis B virus, hepatitis C virus; (B) human T lymphotropic virus type 1, Kaposi's sarcoma-associated herpesvirus; (C) Epstein-Barr virus and human papilloma virus. (Figure was adapted with permission from http://www.genome.jp/kegg-bin/show_pathway?hsa05203 (Kanehisa, 2013))

5.2 Colorectal carcinogenesis

The molecular pathogenesis of colorectal cancer has been mapped out and has been focusing on the molecular players like *APC*, *K-ras*, *TP53*, *DCC*, *SMAD2* and *SMAD4* (Fearon and Vogelstein, 1990, Arnold et al., 2005, Fearon, 2011, Kheirelseid et al., 2013, Remo et al., 2012). Nevertheless, there is also literature indicating that chronic inflammation plays a role in colorectal carcinogenesis (Balkwill and Mantovani, 2001, Coussens and Werb, 2002, Koehne and Dubois, 2004, Galon et al., 2006, Mantovani et al., 2008, Kamp et al., 2011, Vendramini-Costa and Carvalho, 2012, Itzkowitz and Yio, 2004, Greten et al., 2004). Interestingly, viral-mediated colorectal carcinogenesis has been proposed (Enam et al., 2002, Niv et al., 2005, Collins et al., 2011) and many studies have been conducted to demonstrate the pathological association (Chiaravalli et al., 2013, Mou et al., 2012, Vilkin and Niv, 2011, Niv et al., 2010, Lin et al., 2008, Theodoropoulos et al., 2005).

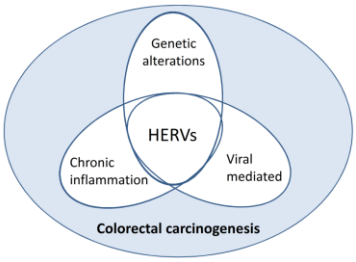


Figure 5.2. A proposed relationship of HERVs in colorectal carcinogenesis. Ancient germ cell infections by exogenous retroviruses has led to the families of HERVs been integrated into the human genome. The integration of retroviral sequences has resulted in insertional polymorphisms (and/or mutagenesis). Hypomethylation of endogenous long terminal repeat (LTR) was recently demonstrated to play a contributory role in carcinogenesis. The Gag, Pol and Env proteins of HERVs are antigenic, and are capable of inducing immune responses. With persistent expression, chronic inflammation may ensue. Given that HERVs are of viral origin and are capable of mediating genetic alterations and chronic inflammation, HERVs present themselves as a putative player in colorectal carcinogenesis.

Taking all into account, the human endogenous retroviruses (HERVs) present themselves as a likely player in the development of colorectal cancer (Figure 5.2), given that HERVs are of viral origin, and have been associated with numerous cancers (Mullins and Linnebacher, 2012, Cegolon et al., 2013) and inflammatory conditions (Reynier et al., 2009, Manghera and Douville, 2013).

5.3 The causal role of HERV-H in colorectal carcinogenesis

As previously mentioned in the preceding chapters, the aetiopathogenic role of HERVs in cancer development is not well understood. Although many studies have detected increased HERV expression in cancers (Dolei, 2006, Cegolon et al., 2013), only HERV-K viral proteins Rec and Np9 have been proven to be oncogenic via the derepression of c-*myc* (Denne et al., 2007). Clinical findings have shown that HERV-H is preferentially expressed in colorectal cancer but not in normal tissue (Liang et al., 2007b, Alves et al., 2008, Wentzensen et al., 2007, Liang et al., 2009). This has led to the current study which aims to unravel the role of HERV-H in colorectal carcinogenesis as colorectal cancer is the most common cancer in both men and women in Singapore (Wong and Eu, 2007, Teo and Soo, 2013). Here, for the first time, our study has demonstrated the oncogenic involvement of HERV-H in colorectal cancer cells. A high distribution frequency (89.9%) of HERV-H in sampled Singapore population is concordant with a high age-standardized incidence rate of colorectal cancer in males (39.9 per 100 000) and females

(28.3 per 100 000), suggesting a role of HERV-H in colorectal cancer in Singapore.

The effects of HERV-H overexpression on colorectal cancer cells were evaluated using established *in vitro* models. Elevated HERV-H expression resulted in the augmentation of proliferative activity, migratory capacity, ability to form cancer stem cell-like colonospheres, and expression levels of cancer stem cell markers, CD133 and CD44. In addition, the gain of decreased serum dependence was most noticeable. Furthermore, it was shown that knock-down of HERV-H significantly lowered cancer cell proliferation, substantiating the oncogenic role of HERV-H in the transforming process (Chapter 3). Interestingly, results from real-time PCR array revealed the involvement of MAPK, Wnt and p53 signalling pathways (Chapter 4).

Clearly, the oncogenic role of HERV-H was illuminated with reference to the hallmarks of cancer (Hanahan and Weinberg, 2000), namely, (i) limitless replicative potential, (ii) tissue invasion & metastasis, (iii) self-sufficiency in growth signals and (iv) the deregulation of cellular energetic in Chapter 3. In addition, the signalling pathways that were implicated in HERV-H-mediated transforming process were also consistent with the hallmarks of cancer, such as (i) sustaining proliferative signalling and (ii) tumour-promoting inflammation (Hanahan and Weinberg, 2011) (Chapter 4). In this chapter, in addition to the hallmarks of cancer, the role of HERV-H in the development of colorectal cancer will be further discussed with reference to a

distinct subclass of neoplastic cells – cancer stem cells (Lobo et al., 2007), which are known to play a contributory role in the next generation of cancer hallmarks (Hanahan and Weinberg, 2011).

5.3.1 Endowment of cancer stem cell potency

The recent reports of a link between increased HERV expressions and stem cell potency are particularly interesting. Studies have shown that the transcriptional activity of endogenous retroviral elements is able to regulate the stem cell potency in mouse embryonic stem cells (Macfarlan et al., 2012), and the expression of HERV-H is able to contribute to the pluripotency in human embryonic stem cells (Santoni et al., 2012). A more recent study reported that HERV-K (HML-2) RNA and protein is a marker for human embryonic and induced pluripotent stem cells (Fuchs et al., 2013). In line with these findings, our finding showed HERV-H up-regulation is capable of endowing colorectal cancer cells with cancer stem cell-like characteristics like increased expression of CD133 and CD44 (markers of cancer stem cells) and the ability to form cancer stem cell-like colonospheres (Figures 3.7, 3.14, 3.15, 3.17 and 3.18).

5.3.2 Activated Wnt signalling linked to CD133 in cancer stem cells

CD133 is a widely used marker of cancer stem cells (O'Brien et al., 2007) and it is also regarded as a target gene of the Wnt/β-catenin signalling pathway (Horst et al., 2009a, Katoh and Katoh, 2007). A previous report has

illustrated this by demonstrating the finding of higher Wnt activity in CD133-positive DLD1 cells when compared to CD133-negative DLD1 cells (Deng et al., 2010). In addition, Wnt/β-catenin signalling plays a pivotal role in the growth and maintenance of colonospheres (Kanwar et al., 2010). In line with colonospheres displaying the characteristics of a cancer stem cell, our study demonstrated that HERV-H transfectants showed higher expression of CD133 and Wnt, as well as the ability to form colonospheres when compared to their control cells (Figures 3.7, 3.14, 3.17 and 4.3).

Viewing from another perspective, there are also numerous studies showing the role of Wnt signalling in the preservation of stem cell proliferation and pluripotency (Schepers and Clevers, 2012, Vijayaragavan et al., 2009). Interestingly, signalling pathways that are known to be tumour-promoting, like the Wnt, Notch and Sonic Hedgehog signalling pathways, are also implicated in normal stem cell self-renewal in the same tissue (Cheng et al., 2011). This implies that Wnt signalling is also playing an important role in the malignancy transformation of embryonic stem cells. In line with these findings, our finding has showed that as HERV-H transfected colorectal cancer cells were endowed with cancer stem cells characteristics, Wnt signalling pathway was implicated (Figures 3.7, 3.14, 3.17 and 4.3).

It is relevant to note that the activation of Wnt pathway can result in the elevation of DKK1 (inhibitor) expression via its negative feedback regulation (Niida et al., 2004). This cause and effect phenomenon was

observed in our study, indicating that the activated Wnt pathway had indeed upregulated the expression of DKK1 (Figure 4.3).

5.3.3 Activated MAPK/ERK signalling linked to CD133 and CD44

CD133 and CD44 are widely used marker of cancer stem cells (O'Brien et al., 2007, Dalerba et al., 2007). CD133 was found to significantly activate MAPK/ERK pathway in U87MG human glioblastoma cells (Dong et al., 2010) and CD133 positive primary colon cancer cells (Wang et al., 2010). Upregulation of CD44 variants was also shown to be linked to MAPK/ERK pathway in primary T cells (Weg-Remers et al., 2001, Herishanu et al., 2011). Additionally, the MAPK/ERK and p38 pathway had also been shown to increase CD44 RNA and CD44 alternate splicing, respectively, in prostate cancer cells (Robbins et al., 2008). The activated ERK pathway was also found to induce the upregulation CD44 in fibrosarcoma, bladder carcinoma and rhabdomyosarcoma (Tanimura et al., 2003). Consistent with these findings, our study demonstrated that HERV-H up-regulation is capable of increasing the expression levels of CD133 and CD44 while activating the MAPK/ERK signalling pathway (Figures 3.14, 3.15, 3.17, 3.18, 4.3 and 4.8).

5.3.4 EGR1 dictates cellular migratory capability

It appears that EGR1 plays a central role in the HERV-H mediated transforming process (Figure 5.3). EGR1, a nuclear phosphoprotein (Milbrandt, 1987) that is inducible by cytokines, growth factors, and stresses like radiation, injury or mechanical stress (Liu et al., 1998, O'Donovan et al.,

1999, Gashler and Sukhatme, 1995), has been shown previously to associate closely with the MAPK signalling pathway. Specifically, increased transcription of EGR1 is mediated by the MAPK signalling pathway, of which all 3 families, p42/44 MAPK, JNK/SAPK and p38 MAPK are involved (Cohen et al., 1996, Harada et al., 1996, Hipskind et al., 1994, Hodge et al., 1998, Lim et al., 1998, Rolli et al., 1999). Given that EGR1 is able to regulate the transcription of heparanase, an endoglycosidase that degrades key components of the extracellular matrix and basement membranes (de Mestre et al., 2005) and that it is also able to serve as a positive regulator of EGF-induced cell migration (Tarcic et al., 2012), EGR1 is playing a part in the regulation of cell migration. The augmented migratory capability and increased expression of EGR1 in HERV-H transfected colorectal cancer cells when compared to control cells is consistent with these previous findings (Figures 3.6 and 4.3).

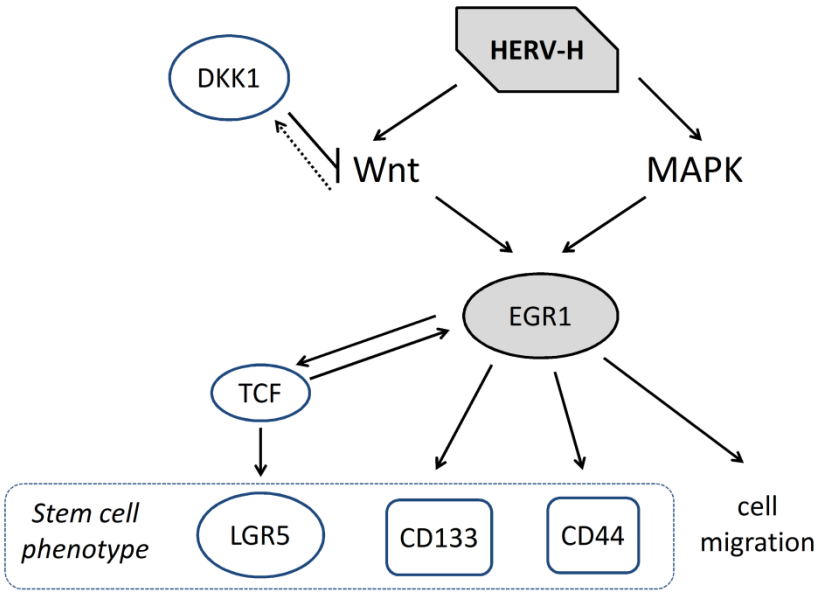


Figure 5.3. Putative central role of EGR1 in HERV-H mediating transforming process.

5.3.5 EGR1 and CD133: direct correlationship

CD133 is a widely used marker of cancer stem cells (O'Brien et al., 2007) . Previous data has shown that EGR1 is highly correlated to CD133 expression (Spearman r = 0.625; P = 0.022; (Ernst et al., 2011)). Consistent with these data, our study demonstrated the concordant expression of CD133 and EGR1 in LS174T:HERV-H_GFP cells (Figures 3.17 and 4.3).

5.3.6 Involvement of EGR1, CD133 and LGR5

In the past 5-6 years, a number of studies had shown that EGR1 may act as a regulator of TCF4 (transcription factor 4) which then acts on LGR5 (a stem cell marker) (Hirsch et al., 2013) via the Wnt/β-catenin pathway (Saegusa et al., 2008, Barker et al., 2009, Barker et al., 2007). Furthermore, it was also demonstrated that aberrant activation of the Wnt pathway increases the LGR5+ stem cell numbers when stimulated by Wnt agonist (Kim et al., 2005). Taken together, this suggests that HERV-H may play a role in the EGR1/CD133/LGR5 network via the Wnt pathway. While this offers a plausible link when cells are endowed with stem cell potency, further studies are needed to determine this network.

5.3.7 Consistency test: role of EGR1 and p53 in proliferation

A concept of two-pulse “consistency test”, which involves EGR1 and p53, was recently proposed (Zwang et al., 2012). The consistency test was put forward to account for the ability of cells to disregard sporadic signals in a

plethora of biological signalling molecules within a cell microenvironment. Specifically, the test requires two consecutive pulses of signalling for mitogenesis. However, it is the second pulse of signalling that instructs cell proliferation, since the second pulse simultaneously downregulates p53-mediated functions which are upregulated by the first pulse of signalling (Figure 5.4) (Zwang et al., 2011). This implies that with the loss of p53 functions, a characteristic of many human cancers, tumour cells would be more responsive to stimulation by very short exposures to proliferative signals. Our finding of increased expression of EGR1 and reduced expression of *TP53* aligns with this novel concept.

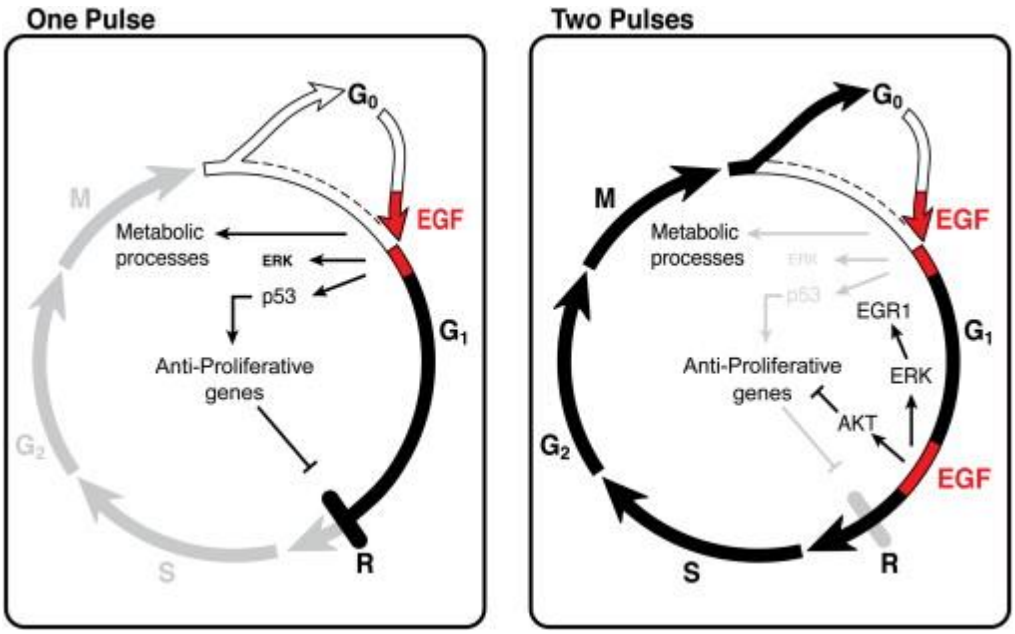


Figure 5.4. Schematic representation of the proposed consistency test of the cell cycle. The first pulse of EGF (epidermal growth factor; a mitogenic signal) enhances lipid metabolism and induces p53 activation. Activation of p53 induces the expression of antiproliferation genes. The second pulse of EGF (mitogenic signal) enhances the activation of ERK/EGR1 and induces PI3K/AKT signalling. This results in a suppressed expression of antiproliferative genes and allows cells to cross the restriction point (R) and enter the S phase (Adapted from (Zwang et al., 2011)).

5.3.8 Activated Wnt activity linked to proliferative and migratory capability

Wnt activity in colorectal cancer stem cells is well established (de Sousa et al., 2011, Vermeulen et al., 2010). Similarly, CD44 is a widely used marker of cancer stem cells (Dalerba et al., 2007). When compared to control cells, HERV-H transfected colorectal cancer cells demonstrated augmented proliferative and migratory capability, and increased expression of CD44 and Wnt11. This is consistent with a previous study that has shown Wnt11 significantly increases the proliferation, migration and invasion activities as well as enhances the expression levels of CD44 in HCT-116 colorectal cancer cells (Nishioka et al., 2013). Interestingly, down-regulation of MAP4K1 is associated with suppressed invasive ability of RKO colon cancer cells (Yang et al., 2006). The decreased expression of MAP4K1 and the lack of gain-of-invasive ability in LS174T:HERV-H_GFP colorectal cancer cells when compared to their control cells are in line with these previous findings.

5.3.9 Conferment of reduced serum dependence and colonospheres-forming ability

The observations of reduced serum dependence and gain of colonospheres-forming ability following HERV-H transduction warrant future investigations. Cancer stem cells, which is considered a principal driving force of tumorigenesis and the initiation of metastases, play an important role in tumour relapse (Nguyen et al., 2012). With phenotypic epithelial-mesenchymal transition (EMT), cancer stem cells metastasize to a distant site

and seed a new tumour, a clinical situation that is associated with poor prognosis. Nevertheless, the key to successful initiation of a new tumour depends not only on the replicative capability of the cancer stem cells but also on the microenvironment of the targeted tissues and the immune system (Medema, 2013). However, should the cancer stem cells be endowed with characteristic like reduced serum dependence, the consequences will be undesirably augmented and unfavourable, since they are able to survive in nutrients-poor sites. In this respect, patients with cancer cachexia may not be spared from tumour relapse if cancer stem cells possess both self-renewal ability and reduced serum dependence characteristic. It is currently not clear what the link between cancer stem cells and the Warburg effect is (Chapter 3). In addition, the link between cancer stem cells quiescence and reduced serum dependence is also unknown, although quiescent cancer stem stems are known to be resistant to cytotoxic chemotherapy and have slow cell cycling states (Li et al., 2008). Further studies are needed to elucidate these relationships.

5.4 Criteria for HERV pathogenicity

First set in 1890, Robert Koch's postulates are widely regarded as the gold standard for establishing the microbiological aetiology of infection and disease (Table 5.2) (Walker et al., 2006). Two centuries later, in a period of time when advances in biotechnology are making impactful contributions to medical knowledge, a broader framework is necessary to identify a pathogenic agent in an intricate host-pathogen relationship (Nelson et al., 2012, Fredericks and Relman, 1996, Segre, 2013). A set of nine criteria were first

proposed by Austin Bradford Hill and put forward to evaluate a putative factor for causality (Table 5.3) (Hill, 1965). While this set of criteria is now widely accepted (Sarid and Gao, 2011), a simplified version had been recently suggested to consider the potential role of HERV in disease (Young et al., 2013).

Table 5.2. Koch's postulates (Walker et al., 2006)

1	The organism must be shown to be invariably present in characteristic form and arrangement in the diseased tissue.
2	The organism, which from its relationship to the diseased tissue appears to be responsible for the disease, must be isolated and grown in pure culture.
3	The pure culture must be shown to induce the disease experimentally.
4	The organism should be re-isolated from the experimentally infected subject.

Table 5.3. Hill's criteria for causation (Hill, 1965, Sarid and Gao, 2011, Carrillo-Infante et al., 2007, Voisset et al., 2008)

Criteria		**Description**
1	Strength of association	Association between an agent and disease
2	Consistency	Consistency of disease association, i.e. consistent findings across studies
3	Specificity	Specificity of the association, i.e. infection is unique to exposure of agent

4	Temporality	Infection precedes disease onset
5	Biological gradient	Dose-response, i.e. an increased dose of agent induces more rapid onset or more severe conditions
6	Biological plausibility	Does the causal relationship seem reasonable when related to current knowledge?
7	Coherence	Is the relationship in line with present knowledge of the disease?
8	Experimental evidence	Does experimental intervention affect disease outcome?

Using Young's HERV criteria for causation, our studies showed that HERV-H is almost able to satisfying all four prescribed criteria (Table 5.4), indicating a close fit to the definition of HERV pathogenicity in colorectal carcinogenesis. Briefly, the four met criteria are (1) the ***discovery*** of selective HERV-H expression in colorectal cancers but not in normal colonic tissues, (2) the ***identification*** of HERV-H being the individual provirus involved in colorectal cancers, (3) the ***description*** of a constructed model for HERV-H involvement, in this case the HERV-H transfected colorectal cancer cell lines, and lastly (4) the ***validation*** carried out for testing the role of HERV-H in colorectal carcinogenesis, which demonstrated the tumorigenic effects of HERV-H and the implicated signalling pathways.

Table 5.4. Four phases of HERV pathogenicity for deliberation. HERV-H is put to assessment using these recent HERV-specific criteria for causality (Young et al., 2013).

Phases		Description	In this study and available literatures
A	**Discovery**	Detection of HERV expression in disease consistently	***Yes***, HERV-H transcripts were detected (Wentzensen et al., 2004, Liang et al., 2007b, Wentzensen et al., 2007, Alves et al., 2008, Liang et al., 2009).
B	**Identification**	Definition of individual proviruses involved	***Yes***, specifically HERV-H is involved. High prevalence rate of HERV-H in sampled Singapore population is concordant with high incidence rate of colorectal cancer in Singapore. (**Results in Chapter 2**)
C	**Description**	Construction of a specific model for HERV involvement	***Yes.*** Construction of HERV-H transfected colorectal cancer cell lines done. (**Results in Chapter 3**)
D	**Validation**	Testing to prove/disprove model	***Yes***. HERV-H transfected colorectal cancer cell lines showed oncogenic potential of HERV-H peptide. Implicated signalling pathways were made known by real time PCR array results. (**Results in Chapters 3 & 4**)

5.5 Conclusion

An aetiopathogenic role of HERVs in cancer development has long been postulated, but direct molecular evidence of their involvement in disease has been incomplete. The studies reported in this thesis demonstrated that

HERV-H is highly prevalent (89.3%) in the sampled population of Singapore where its incidence rate of colorectal cancer remains consistently high. Given the findings that HERV-H overexpression endowed colorectal cancer cells with greater cell proliferation, motility, ability to form cancer stem cell-like colonospheres and expression of stem cell markers (CD133 and CD44), as well as reduced serum dependence, it is reasonable to ascribe the oncogenic potential of HERV-H for colorectal carcinogenesis. An additional level of evidence would come from the finding that knock-down of HERV-H significantly reduced cancer cell proliferation. Using real-time PCR array, mitogen-activated protein kinases (MAPK), Wnt and p53 pathways were all found to be implicated in HERV-H-mediated oncogenic signalling. This information provided an invaluable resource for rationalizing the therapeutic approach to treat colorectal cancers. Together with previous clinical reports, our data demonstrates the tumorigenic role of HERV-H in colorectal cancer and highlights the molecular basis underlying its involvement. Importantly, the work described in this thesis fits into the Young's four-phase framework of HERV pathogenicity, illuminating the causality of HERV-H in pathogenesis of colorectal cancer in a structured, phasic approach.

Future directions

Chapter 6

CHAPTER 6

FUTURE DIRECTIONS

6.1 Proposed future areas of work

This study has provided numerous insights into the possible role of HERV-H in the pathogenesis of colorectal cancer. The present findings, which probe into HERV-H oncogenic characteristics and implicated signalling pathways, is to our knowledge without precedence. The unravelling of signalling pathways that involved Wnt, MAPK and p53 open up therapeutic vistas for HERV-H implicated colorectal cancers.

However, further works are necessary to probe for the precise molecular target in the light of different signalling pathways are involved. In addition, it is also helpful to determine the consequences of HERV-H expression in colorectal cancer. This may aid in the selection of aggressive versus conventional therapeutic interventions in different clinical settings. Hence, the future areas of research would be to further identify molecular events that are important to HERV-H mediating transformation process as well as to demonstrate the prognostic value of HERV-H expression in colorectal cancer. Scope for future works includes the following:

6.1.1 Characterisation of aberrant initiating points in signalling pathways

In this study, it was found that Wnt, MAPK and p53 signalling pathways are involved in the HERV-H mediating transformation process. To

enhance the translational value of this knowledge, it is worthwhile to investigate the precise molecular target in the implicated signalling pathways. This can be done using a specific inhibitor of the pathway or inhibitors of each individual pathway at various intermediate steps. The resultant information will provide invaluable resources for the development of novel molecular therapeutic interventions. In addition, the interacting partner(s) of HERV-H can be identified by pull-down assay, co-immunoprecipitation (Co-IP) or yeast two hybrid screening. With proteomic identification of the interacting partner proteins, the role of HERV-H in colorectal carcinogenesis can be further defined and better molecular targeted cancer therapeutic drugs can be devised.

6.1.2 *In vivo* discovery of HERV-H tumorigenicity

The tumorigenicity of HERV-H transfected colorectal cancer cells can be examined by using animal models, for instance, the use of nude mice (Sharkey and Fogh, 1984). Mice with HERV-H transfected tumours can also be further employed in another setting where they can be treated with different pharmacological agents to evaluate how these tumours regress. With this testing system, the tumour-killing efficiency of various pharmacological agents can be exploited efficiently.

6.1.3 Implications of HERV-H sequence homology

A high level of HERV-H sequence conservation was observed in this study (Chapter 2). A thorough evaluation of the biological significance of this

evolutionarily conserved region is warranted as highly conserved DNA sequences are conventionally thought to possess functional value. Using a variety of bioinformatics resources, studies may involve functional *in silico* analysis of conserved gene segment (Nimrod et al., 2005), assignment of function to conserved noncoding sequences (Meisler, 2001) and comparative analysis of noncoding regions across vertebrate genomes (Venkatesh et al., 2006, Woolfe et al., 2005).

6.1.4 The significance of reduced serum dependence

The discovery of reduced serum dependence following HERV-H transduction is remarkably intriguing as it suggests cancer resilience that associates well with poor clinical outcome. However, the molecular mechanism underlying this characteristic remains to be elucidated.

It is suggested that an altered metabolism of cancer cells, the Warburg effect, is conducive to cell proliferation (Vander Heiden et al., 2009). The link between reduced serum dependence and the Warburg effect is unknown, although both are enabling cancer cells to adapt well in a nutrients-deficient environment. Thus, it is interesting to examine the effect of HERV-H up-regulation on the Warburg effect in colorectal cancer cells. In addition, the clinical implications of reduced serum dependence can be tested out using animal models. It is anticipated that the effect of reduced serum dependence will associate with acute fulminant metastases and poor survival rate.

6.1.5 The role of HERV-H in the EGR1/CD133/LGR5 network

Heightened expression of EGR1 and CD133 is associated with cancer stem cells potency (Ernst et al., 2011). LGR5 is another stem cell marker which is used in colorectal cancer studies (Hirsch et al., 2013, Barker et al., 2007) and is implicated in the EGR1/CD133/LGR5 network (Figure 5.3) (Ernst et al., 2011) . Our study has shown that HERV-H overexpression endows colorectal cancer cells with cancer stem cell-like characteristics like increased expression of EGR1 and CD133, as well as colonospheres-forming ability. It is interesting to know whether LGR5 is also implicated in the HERV-H mediated transformation process. Further studies involving overexpression and knockdown assays of EGR1 and LGR5 will provide insights into the molecular mechanisms underlying this pathway. In addition, to examine how HERV-H mediated transformation process is intertwined with stem cell pluripotency, the expression profile of embryonic stem cells markers like SOX2, OCT4 and Nanog (Mitsui et al., 2003, Pan and Thomson, 2007, Okumura-Nakanishi et al., 2005, Niwa et al., 2000) can be determined.

6.1.6 Evaluation of prognostic value of HERV-H in Singapore cohort of colorectal cancer patients

Clinical studies demonstrating HERV-H expression in colorectal cancers are already available (Liang et al., 2007b, Alves et al., 2008, Wentzensen et al., 2007, Liang et al., 2009). To demonstrate the significance of HERV-H expression in the Singapore cohort of colorectal cancer patients, a prospective cohort study or a retrospective case-control study can be initiated.

A large-scale study is needed for clinical relevance. Sampling should include cohorts of colorectal cancer patients and healthy individuals. A clearly defined biomarker derived from this large-scale cohort study would serve well as a stratification factor for clinical outcome. This will aid in predicting therapeutic options and colorectal cancer survival in the Singapore cohort of patients.

References

Chapter 7

ABDULKADIR, S. A., QU, Z., GARABEDIAN, E., SONG, S. K., PETERS, T. J., SVAREN, J., CARBONE, J. M., NAUGHTON, C. K., CATALONA, W. J., ACKERMAN, J. J., GORDON, J. I., HUMPHREY, P. A. & MILBRANDT, J. 2001. Impaired prostate tumorigenesis in Egr1-deficient mice. *Nat Med,* 7**,** 101-7.

AGATHEESWARAN, S., SINGH, S., BISWAS, S., BISWAS, G., CHANDRA PATTNAYAK, N. & CHAKRABORTY, S. 2013. BCR-ABL mediated repression of miR-223 results in the activation of MEF2C and PTBP2 in chronic myeloid leukemia. *Leukemia,* 27**,** 1578-80.

AGGARWAL, B. B., SHISHODIA, S., SANDUR, S. K., PANDEY, M. K. & SETHI, G. 2006. Inflammation and cancer: how hot is the link? *Biochem Pharmacol,* 72**,** 1605-21.

AGONI, L., GOLDEN, A., GUHA, C. & LENZ, J. 2012. Neandertal and Denisovan retroviruses. *Curr Biol,* 22**,** R437-8.

ALVES, P. M., LEVY, N., STEVENSON, B. J., BOUZOURENE, H., THEILER, G., BRICARD, G., VIATTE, S., AYYOUB, M., VUILLEUMIER, H., GIVEL, J. C., RIMOLDI, D., SPEISER, D. E., JONGENEEL, C. V., ROMERO, P. J. & LEVY, F. 2008. Identification of tumor-associated antigens by large-scale analysis of genes expressed in human colorectal cancer. *Cancer Immun,* 8**,** 11.

AMERICAN CANCER SOCIETY. 2012. *Colorectal Cancer: What are the risk factors for colorectal cancer?* [Online]. American Cancer Society, Inc. Available: http://www.cancer.org/cancer/colonandrectumcancer/detailedguide/colorectal-cancer-risk-factors [Accessed 25 January 2013 2013].

ANDERSSON, M. L., LINDESKOG, M., MEDSTRAND, P., WESTLEY, B., MAY, F. & BLOMBERG, J. 1999. Diversity of human endogenous retrovirus class II-like sequences. *J Gen Virol,* 80 (Pt 1)**,** 255-60.

ANGERS, S. & MOON, R. T. 2009. Proximal events in Wnt signal transduction. *Nat Rev Mol Cell Biol,* 10**,** 468-77.

ANTONY, J. M., DESLAURIERS, A. M., BHAT, R. K., ELLESTAD, K. K. & POWER, C. 2011. Human endogenous retroviruses and multiple sclerosis: innocent bystanders or disease determinants? *Biochim Biophys Acta,* 1812**,** 162-76.

ARIKAWA, E., SUN, Y., WANG, J., ZHOU, Q., NING, B., DIAL, S. L., GUO, L. & YANG, J. 2008. Cross-platform comparison of SYBR Green real-time PCR with TaqMan PCR, microarrays and other gene expression measurement technologies evaluated in the MicroArray Quality Control (MAQC) study. *BMC Genomics,* 9**,** 328.

ARMBRUESTER, V., SAUTER, M., ROEMER, K., BEST, B., HAHN, S., NTY, A., SCHMID, A., PHILIPP, S., MUELLER, A. & MUELLER-LANTZSCH, N. 2004. Np9 protein of human endogenous retrovirus K interacts with ligand of numb protein X. *J Virol,* 78**,** 10310-9.

ARNOLD, C. N., GOEL, A., BLUM, H. E. & BOLAND, C. R. 2005. Molecular pathogenesis of colorectal cancer: implications for molecular diagnosis. *Cancer,* 104**,** 2035-47.

BABU, S. N., CHETAL, G. & KUMAR, S. 2012. Macrophage migration inhibitory factor: a potential marker for cancer diagnosis and therapy. *Asian Pac J Cancer Prev,* 13**,** 1737-44.

BADENHOOP, K., DONNER, H., NEUMANN, J., HERWIG, J., KURTH, R., USADEL, K. H. & TONJES, R. R. 1999. IDDM patients neither show humoral reactivities against endogenous retroviral envelope protein nor do they differ in retroviral mRNA expression from healthy relatives or normal individuals. *Diabetes,* 48**,** 215-8.

BAFICO, A., LIU, G., YANIV, A., GAZIT, A. & AARONSON, S. A. 2001. Novel mechanism of Wnt signalling inhibition mediated by Dickkopf-1 interaction with LRP6/Arrow. *Nat Cell Biol,* 3**,** 683-6.

BAI, X., WU, L., LIANG, T., LIU, Z., LI, J., LI, D., XIE, H., YIN, S., YU, J., LIN, Q. & ZHENG, S. 2008. Overexpression of myocyte enhancer factor 2 and histone hyperacetylation in hepatocellular carcinoma. *J Cancer Res Clin Oncol,* 134**,** 83-91.

BAKER, S. J., PREISINGER, A. C., JESSUP, J. M., PARASKEVA, C., MARKOWITZ, S., WILLSON, J. K., HAMILTON, S. & VOGELSTEIN, B. 1990. p53 gene mutations occur in combination with 17p allelic deletions as late events in colorectal tumorigenesis. *Cancer Res,* 50**,** 7717-22.

BALADA, E., VILARDELL-TARRES, M. & ORDI-ROS, J. 2010. Implication of human endogenous retroviruses in the development of autoimmune diseases. *Int Rev Immunol,* 29**,** 351-70.

BALESTRIERI, E., ARPINO, C., MATTEUCCI, C., SORRENTINO, R., PICA, F., ALESSANDRELLI, R., CONIGLIO, A., CURATOLO, P., REZZA, G., MACCIARDI, F., GARACI, E., GAUDI, S. & SINIBALDI-VALLEBONA, P. 2012. HERVs expression in Autism Spectrum Disorders. *PLoS One,* 7**,** e48831.

BALKWILL, F. & MANTOVANI, A. 2001. Inflammation and cancer: back to Virchow? *Lancet,* 357**,** 539-45.

BALTIMORE, D. 1971. Expression of animal virus genomes. *Bacteriol Rev,* 35**,** 235-41.

BANDEA, C. I. 1983. A new theory on the origin and the nature of viruses. *J Theor Biol,* 105**,** 591-602.

BANNERT, N. & KURTH, R. 2004. Retroelements and the human genome: new perspectives on an old relation. *Proc Natl Acad Sci U S A,* 101 Suppl 2**,** 14572-9.

BANNERT, N. & KURTH, R. 2006. The evolutionary dynamics of human endogenous retroviral families. *Annu Rev Genomics Hum Genet,* 7**,** 149-73.

BARBULESCU, M., TURNER, G., SEAMAN, M. I., DEINARD, A. S., KIDD, K. K. & LENZ, J. 1999. Many human endogenous retrovirus K (HERV-K) proviruses are unique to humans. *Curr Biol,* 9**,** 861-8.

BARGONETTI, J., REYNISDOTTIR, I., FRIEDMAN, P. N. & PRIVES, C. 1992. Site-specific binding of wild-type p53 to cellular DNA is inhibited by SV40 T antigen and mutant p53. *Genes Dev,* 6**,** 1886-98.

BARKER, N., RIDGWAY, R. A., VAN ES, J. H., VAN DE WETERING, M., BEGTHEL, H., VAN DEN BORN, M., DANENBERG, E., CLARKE, A. R., SANSOM, O. J. & CLEVERS, H. 2009. Crypt stem cells as the cells-of-origin of intestinal cancer. *Nature,* 457**,** 608-11.

BARKER, N., VAN ES, J. H., KUIPERS, J., KUJALA, P., VAN DEN BORN, M., COZIJNSEN, M., HAEGEBARTH, A., KORVING, J., BEGTHEL, H., PETERS, P. J. & CLEVERS, H. 2007. Identification of stem cells in small intestine and colon by marker gene Lgr5. *Nature,* 449**,** 1003-7.

BARR, J. J., AURO, R., FURLAN, M., WHITESON, K. L., ERB, M. L., POGLIANO, J., STOTLAND, A., WOLKOWICZ, R., CUTTING, A. S., DORAN, K. S., SALAMON, P., YOULE, M. & ROHWER, F. 2013. Bacteriophage adhering to mucus provide a non-host-derived immunity. *Proc Natl Acad Sci U S A*.

BASU, M. & ROY, S. S. 2013. Wnt/beta-catenin pathway is regulated by PITX2 homeodomain protein and thus contributes to the proliferation of human ovarian adenocarcinoma cell, SKOV-3. *J Biol Chem,* 288**,** 4355-67.

BAUMANN, J. G. 2006. Intracellular restriction factors in mammalian cells--An ancient defense system finds a modern foe. *Curr HIV Res,* 4**,** 141-68.

BELSHAW, R., DAWSON, A. L., WOOLVEN-ALLEN, J., REDDING, J., BURT, A. & TRISTEM, M. 2005. Genomewide screening reveals high levels of insertional polymorphism in the human endogenous retrovirus family HERV-K(HML2): implications for present-day activity. *J Virol,* 79**,** 12507-14.

BELSHE, R. B. 2005. The origins of pandemic influenza--lessons from the 1918 virus. *N Engl J Med,* 353**,** 2209-11.

BENIT, L., DESSEN, P. & HEIDMANN, T. 2001. Identification, phylogeny, and evolution of retroviral elements based on their envelope genes. *J Virol,* 75**,** 11709-19.

BENIT, L., LALLEMAND, J. B., CASELLA, J. F., PHILIPPE, H. & HEIDMANN, T. 1999. ERV-L elements: a family of endogenous retrovirus-like elements active throughout the evolution of mammals. *J Virol,* 73**,** 3301-8.

BERGH, O., BORSHEIM, K. Y., BRATBAK, G. & HELDAL, M. 1989. High abundance of viruses found in aquatic environments. *Nature,* 340**,** 467-8.

BERGOLD, G. H. 1953. On the nomenclature and classification of insect viruses. *Ann N Y Acad Sci,* 56**,** 495-516.

BISHOP, J. M. 1978. Retroviruses. *Annu Rev Biochem,* 47**,** 35-88.

BISTER, K. & DUESBERG, P. H. 1979. Structure and specific sequences of avian erythroblastosis virus RNA: evidence for multiple classes of transforming genes among avian tumor viruses. *Proc Natl Acad Sci U S A,* 76**,** 5023-7.

BLACK, B. L. & OLSON, E. N. 1998. Transcriptional control of muscle development by myocyte enhancer factor-2 (MEF2) proteins. *Annu Rev Cell Dev Biol,* 14**,** 167-96.

BLAISE, S., DE PARSEVAL, N., BENIT, L. & HEIDMANN, T. 2003. Genomewide screening for fusogenic human endogenous retrovirus envelopes identifies syncytin 2, a gene conserved on primate evolution. *Proc Natl Acad Sci U S A,* 100**,** 13013-8.

BLASCHKE, V., REICH, K., MIDDEL, P., LETSCHERT, M., SACHSE, F., HARWIX, S. & NEUMANN, C. 1999. Expression of the CD4+ cell-specific chemoattractant interleukin-16 in mycosis fungoides. *J Invest Dermatol,* 113**,** 658-63.

BLIKSTAD, V., BENACHENHOU, F., SPERBER, G. O. & BLOMBERG, J. 2008. Evolution of human endogenous retroviral sequences: a conceptual account. *Cell Mol Life Sci,* 65**,** 3348-65.

BLOMBERG, J., BENACHENHOU, F., BLIKSTAD, V., SPERBER, G. & MAYER, J. 2009. Classification and nomenclature of endogenous retroviral sequences (ERVs): problems and recommendations. *Gene,* 448**,** 115-23.

BOCK, M. & STOYE, J. P. 2000. Endogenous retroviruses and the human germline. *Curr Opin Genet Dev,* 10**,** 651-5.

BOESE, A., SAUTER, M., GALLI, U., BEST, B., HERBST, H., MAYER, J., KREMMER, E., ROEMER, K. & MUELLER-LANTZSCH, N. 2000. Human endogenous retrovirus protein cORF supports cell transformation and associates with the promyelocytic leukemia zinc finger protein. *Oncogene,* 19**,** 4328-36.

BOULESTEIX, M. & BIEMONT, C. 2005. Transposable elements in mosquitoes. *Cytogenet Genome Res,* 110**,** 500-9.

BRAUN, F., BERTIN-CIFTCI, J., GALLOUET, A. S., MILLOUR, J. & JUIN, P. 2011. Serum-nutrient starvation induces cell death mediated by Bax and Puma that is counteracted by p21 and unmasked by Bcl-x(L) inhibition. *PLoS One,* 6**,** e23577.

BRAY, F., REN, J. S., MASUYER, E. & FERLAY, J. 2013. Global estimates of cancer prevalence for 27 sites in the adult population in 2008. *Int J Cancer,* 132**,** 1133-45.

BREMAN, J. G. & HENDERSON, D. A. 2002. Diagnosis and management of smallpox. *N Engl J Med,* 346**,** 1300-8.

BROWN, L., ONGUSAHA, P. P., KIM, H. G., NUTI, S., MANDINOVA, A., LEE, J. W., KHOSRAVI-FAR, R., AARONSON, S. A. & LEE, S. W. 2007. CDIP, a novel pro-apoptotic gene, regulates TNFalpha-mediated apoptosis in a p53-dependent manner. *EMBO J,* 26**,** 3410-22.

BRUGAROLAS, J., CHANDRASEKARAN, C., GORDON, J. I., BEACH, D., JACKS, T. & HANNON, G. J. 1995. Radiation-induced cell cycle arrest compromised by p21 deficiency. *Nature,* 377**,** 552-7.

BRYANT, M. L., SHERR, C. J., SEN, A. & TODARO, G. J. 1978. Molecular diversity among five different endogenous primate retroviruses. *J Virol,* 28**,** 300-13.

BURGESS, A. W., FAUX, M. C., LAYTON, M. J. & RAMSAY, R. G. 2011. Wnt signaling and colon tumorigenesis--a view from the periphery. *Exp Cell Res,* 317**,** 2748-58.

BUTEL, J. S. 2000. Viral carcinogenesis: revelation of molecular mechanisms and etiology of human disease. *Carcinogenesis,* 21**,** 405-26.

BUTLER, L. M., WANG, R., KOH, W. P. & YU, M. C. 2008. Prospective study of dietary patterns and colorectal cancer among Singapore Chinese. *Br J Cancer,* 99**,** 1511-6.

CALANDRA, T. & ROGER, T. 2003. Macrophage migration inhibitory factor: a regulator of innate immunity. *Nat Rev Immunol,* 3**,** 791-800.

CALCAGNO, D. Q., FREITAS, V. M., LEAL, M. F., DE SOUZA, C. R., DEMACHKI, S., MONTENEGRO, R., ASSUMPCAO, P. P., KHAYAT, A. S., SMITH MDE, A., DOS SANTOS, A. K. & BURBANO, R. R. 2013. MYC, FBXW7 and TP53 copy number variation and expression in Gastric Cancer. *BMC Gastroenterol,* 13**,** 141.

CAPPELL, M. S. 2008. Pathophysiology, clinical presentation, and management of colon cancer. *Gastroenterol Clin North Am,* 37**,** 1-24, v.

CARDELLI, M. & MARCHEGIANI, F. 2012. Good, Bad, Mobile Elements: Genomes Most Successful "Parasites" As Emerging Players in Cell and Organismal Aging. *Curr Pharm Des*.

CARR, C. E., ROWEDDER, H., LUI, C. S., ZLATKOVSKY, I., PAPALIAS, C. W., BOLANDER, J., MYERS, J. W., BUSTILLO, J., ROTHBERG, J. M., ZUBER, M. T. & RUVKUN, G. 2013. Radiation resistance of sequencing chips for in situ life detection. *Astrobiology,* 13**,** 560-9.

CARRILLO-INFANTE, C., ABBADESSA, G., BAGELLA, L. & GIORDANO, A. 2007. Viral infections as a cause of cancer (review). *Int J Oncol,* 30**,** 1521-8.

CAVAZZA, A., MOIANI, A. & MAVILIO, F. 2013. Mechanisms of retroviral integration and mutagenesis. *Hum Gene Ther,* 24**,** 119-31.

CEGOLON, L., SALATA, C., WEIDERPASS, E., VINEIS, P., PALU, G. & MASTRANGELO, G. 2013. Human endogenous retroviruses and cancer prevention: evidence and prospects. *BMC Cancer,* 13**,** 4.

CENTER, D. M. & CRUIKSHANK, W. 1982. Modulation of lymphocyte migration by human lymphokines. I. Identification and characterization of chemoattractant activity for lymphocytes from mitogen-stimulated mononuclear cells. *J Immunol,* 128**,** 2563-8.

CENTERS FOR DISEASE CONTROL AND PREVENTION. 2011. *Colorectal Cancer Risk Factors* [Online]. Atlanta, USA: Centers for Disease Control and Prevention Available: http://www.cdc.gov/cancer/colorectal/basic_info/risk_factors.htm [Accessed 25 January 2013 2013].

CHA, S. W., TADJUIDJE, E., TAO, Q., WYLIE, C. & HEASMAN, J. 2008. Wnt5a and Wnt11 interact in a maternal Dkk1-regulated fashion to activate both canonical and non-canonical signaling in Xenopus axis formation. *Development,* 135**,** 3719-29.

CHEAH, P. Y. 2009. Recent advances in colorectal cancer genetics and diagnostics. *Crit Rev Oncol Hematol,* 69**,** 45-55.

CHEN, T., MENG, Z., GAN, Y., WANG, X., GU, Y., XU, X., TANG, J., ZHOU, H., ZHANG, X., GAN, X., VAN NESS, C., XU, F., XU, G., HUANG, L., ZHANG, X., FANG, Y., WU, J., ZHENG, S., JIN, J., HUANG, W. & XU, R. 2013. The viral oncogene Np9 acts as a critical molecular switch for co-activating beta-catenin, ERK, Akt and Notch1 and promoting the growth of human leukemia stem/progenitor cells. *Leukemia*.

CHENG, L., ALEXANDER, R., ZHANG, S., PAN, C. X., MACLENNAN, G. T., LOPEZ-BELTRAN, A. & MONTIRONI, R. 2011. The clinical and therapeutic implications of cancer stem cell biology. *Expert Rev Anticancer Ther,* 11**,** 1131-43.

CHIARAVALLI, A. M., LONGHI, E., VIGETTI, D., DE STEFANO, F. I., DELEONIBUS, S., CAPELLA, C., SOLCIA, E. & PARRAVICINI, C. 2013. Gastrointestinal cancers reactive for the PAb416 antibody against JCV/SV40 T-Ag lack JCV DNA sequences while showing a distinctive pathologic profile. *J Clin Pathol,* 66**,** 44-9.

CHO, H. H., KYOUNG, K. M., SEO, M. J., KIM, Y. J., BAE, Y. C. & JUNG, J. S. 2006. Overexpression of CXCR4 increases migration and proliferation of human adipose tissue stromal cells. *Stem Cells Dev,* 15**,** 853-64.

CHO, S. H., PARK, Y. S., KIM, H. J., KIM, C. H., LIM, S. W., HUH, J. W., LEE, J. H. & KIM, H. R. 2012. CD44 enhances the epithelial-mesenchymal transition in association with colon cancer invasion. *Int J Oncol,* 41**,** 211-8.

CHRISTENSEN, T. 2010. HERVs in neuropathogenesis. *J Neuroimmune Pharmacol,* 5**,** 326-35.

COHEN, D. M., GULLANS, S. R. & CHIN, W. W. 1996. Urea inducibility of egr-1 in murine inner medullary collecting duct cells is mediated by the serum response element and adjacent Ets motifs. *J Biol Chem,* 271**,** 12903-8.

COHEN, E. D., TIAN, Y. & MORRISEY, E. E. 2008. Wnt signaling: an essential regulator of cardiovascular differentiation, morphogenesis and progenitor self-renewal. *Development,* 135**,** 789-98.

COHEN, M., POWERS, M., O'CONNELL, C. & KATO, N. 1985. The nucleotide sequence of the env gene from the human provirus ERV3 and isolation and characterization of an ERV3-specific cDNA. *Virology,* 147**,** 449-58.

COLLINS, A. T., BERRY, P. A., HYDE, C., STOWER, M. J. & MAITLAND, N. J. 2005. Prospective identification of tumorigenic prostate cancer stem cells. *Cancer Res,* 65**,** 10946-51.

COLLINS, D., HOGAN, A. M. & WINTER, D. C. 2011. Microbial and viral pathogens in colorectal cancer. *Lancet Oncol,* 12**,** 504-12.

COLLOT-TEIXEIRA, S., BASS, J., DENIS, F. & RANGER-ROGEZ, S. 2004. Human tumor suppressor p53 and DNA viruses. *Rev Med Virol,* 14**,** 301-19.

COLOTTA, F., ALLAVENA, P., SICA, A., GARLANDA, C. & MANTOVANI, A. 2009. Cancer-related inflammation, the seventh hallmark of cancer: links to genetic instability. *Carcinogenesis,* 30**,** 1073-81.

CONRAD, B., WEISSMAHR, R. N., BONI, J., ARCARI, R., SCHUPBACH, J. & MACH, B. 1997. A human endogenous retroviral superantigen as candidate autoimmune gene in type I diabetes. *Cell,* 90**,** 303-13.

CONROY, H., MAWHINNEY, L. & DONNELLY, S. C. 2010. Inflammation and cancer: macrophage migration inhibitory factor (MIF)--the potential missing link. *QJM,* 103**,** 831-6.

COOK, L. B., ELEMANS, M., ROWAN, A. G. & ASQUITH, B. 2013. HTLV-1: persistence and pathogenesis. *Virology,* 435**,** 131-40.

CORDAUX, R. & BATZER, M. A. 2009. The impact of retrotransposons on human genome evolution. *Nat Rev Genet,* 10**,** 691-703.

CORDONNIER, A., CASELLA, J. F. & HEIDMANN, T. 1995. Isolation of novel human endogenous retrovirus-like elements with foamy virus-related pol sequence. *J Virol,* 69**,** 5890-7.

COUSSENS, L. M. & WERB, Z. 2002. Inflammation and cancer. *Nature,* 420**,** 860-7.

CRABTREE, G. R. & OLSON, E. N. 2002. NFAT signaling: choreographing the social lives of cells. *Cell,* 109 Suppl**,** S67-79.

CRIGHTON, D., WILKINSON, S., O'PREY, J., SYED, N., SMITH, P., HARRISON, P. R., GASCO, M., GARRONE, O., CROOK, T. & RYAN, K. M. 2006. DRAM, a p53-induced modulator of autophagy, is critical for apoptosis. *Cell,* 126**,** 121-34.

CROCE, C. M. 2008. Oncogenes and cancer. *N Engl J Med,* 358**,** 502-11.

CRUIKSHANK, W. W., KORNFELD, H. & CENTER, D. M. 2000. Interleukin-16. *J Leukoc Biol,* 67**,** 757-66.

CSIBI, A., FENDT, S. M., LI, C., POULOGIANNIS, G., CHOO, A. Y., CHAPSKI, D. J., JEONG, S. M., DEMPSEY, J. M., PARKHITKO, A., MORRISON, T., HENSKE, E. P., HAIGIS, M. C., CANTLEY, L. C., STEPHANOPOULOS, G., YU, J. & BLENIS, J. 2013. The mTORC1 pathway stimulates glutamine metabolism and cell proliferation by repressing SIRT4. *Cell,* 153**,** 840-54.

CULLEN, B. R. 2006. Role and mechanism of action of the APOBEC3 family of antiretroviral resistance factors. *J Virol,* 80**,** 1067-76.

CUNNINGHAM, D., ATKIN, W., LENZ, H. J., LYNCH, H. T., MINSKY, B., NORDLINGER, B. & STARLING, N. 2010. Colorectal cancer. *Lancet,* 375**,** 1030-47.

DABOUSSI, M. J. & CAPY, P. 2003. Transposable elements in filamentous fungi. *Annu Rev Microbiol,* 57**,** 275-99.

DALERBA, P., DYLLA, S. J., PARK, I. K., LIU, R., WANG, X., CHO, R. W., HOEY, T., GURNEY, A., HUANG, E. H., SIMEONE, D. M., SHELTON, A. A., PARMIANI, G., CASTELLI, C. & CLARKE, M. F. 2007. Phenotypic characterization of human colorectal cancer stem cells. *Proc Natl Acad Sci U S A,* 104**,** 10158-63.

DALLAS, M. R., LIU, G., CHEN, W. C., THOMAS, S. N., WIRTZ, D., HUSO, D. L. & KONSTANTOPOULOS, K. 2012. Divergent roles of CD44 and carcinoembryonic antigen in colon cancer metastasis. *FASEB J,* 26**,** 2648-56.

DANG, C. V. 2009. MYC, microRNAs and glutamine addiction in cancers. *Cell Cycle,* 8**,** 3243-5.

DANGEL, A. W., BAKER, B. J., MENDOZA, A. R. & YU, C. Y. 1995. Complement component C4 gene intron 9 as a phylogenetic marker for primates: long terminal repeats of the endogenous retrovirus ERV-K(C4) are a molecular clock of evolution. *Immunogenetics,* 42**,** 41-52.

DANIEL, C., GERLACH, K., VATH, M., NEURATH, M. F. & WEIGMANN, B. 2013. Nuclear factor of activated T cells-A transcription factor family as critical regulator in lung and colon cancer. *Int J Cancer*.

DE LA HERA, B., VARADE, J., GARCIA-MONTOJO, M., LAMAS, J. R., DE LA ENCARNACION, A., ARROYO, R., FERNANDEZ-GUTIERREZ, B., ALVAREZ-LAFUENTE, R. & URCELAY, E. 2013. Role of the human endogenous retrovirus HERV-K18 in autoimmune disease susceptibility: study in the Spanish population and meta-analysis. *PLoS One,* 8**,** e62090.

DE MARTEL, C., FERLAY, J., FRANCESCHI, S., VIGNAT, J., BRAY, F., FORMAN, D. & PLUMMER, M. 2012. Global burden of cancers attributable to infections in 2008: a review and synthetic analysis. *Lancet Oncol,* 13**,** 607-15.

DE MESTRE, A. M., RAO, S., HORNBY, J. R., SOE-HTWE, T., KHACHIGIAN, L. M. & HULETT, M. D. 2005. Early growth response gene 1 (EGR1) regulates heparanase gene transcription in tumor cells. *J Biol Chem,* 280**,** 35136-47.

DE PARSEVAL, N. & HEIDMANN, T. 1998. Physiological knockout of the envelope gene of the single-copy ERV-3 human endogenous retrovirus in a fraction of the Caucasian population. *J Virol,* 72**,** 3442-5.

DE SOUSA, E. M., VERMEULEN, L., RICHEL, D. & MEDEMA, J. P. 2011. Targeting Wnt signaling in colon cancer stem cells. *Clin Cancer Res,* 17**,** 647-53.

DEININGER, P. L. & BATZER, M. A. 2002. Mammalian retroelements. *Genome Res,* 12**,** 1455-65.

DEMETS, R. 2012. Darwin's contribution to the development of the Panspermia theory. *Astrobiology,* 12**,** 946-50.

DENES, A., LOPEZ-CASTEJON, G. & BROUGH, D. 2012. Caspase-1: is IL-1 just the tip of the ICEberg? *Cell Death Dis,* 3**,** e338.

DENG, C., ZHANG, P., HARPER, J. W., ELLEDGE, S. J. & LEDER, P. 1995. Mice lacking p21CIP1/WAF1 undergo normal development, but are defective in G1 checkpoint control. *Cell,* 82**,** 675-84.

DENG, Y. H., PU, X. X., HUANG, M. J., XIAO, J., ZHOU, J. M., LIN, T. Y. & LIN, E. H. 2010. 5-Fluorouracil upregulates the activity of Wnt signaling pathway in CD133-positive colon cancer stem-like cells. *Chin J Cancer,* 29**,** 810-5.

DENNE, M., SAUTER, M., ARMBRUESTER, V., LICHT, J. D., ROEMER, K. & MUELLER-LANTZSCH, N. 2007. Physical and functional interactions of human endogenous retrovirus proteins Np9 and rec with the promyelocytic leukemia zinc finger protein. *J Virol,* 81**,** 5607-16.

DEPIL, S., ROCHE, C., DUSSART, P. & PRIN, L. 2002. Expression of a human endogenous retrovirus, HERV-K, in the blood cells of leukemia patients. *Leukemia,* 16**,** 254-9.

DHILLON, A. S., HAGAN, S., RATH, O. & KOLCH, W. 2007. MAP kinase signalling pathways in cancer. *Oncogene,* 26**,** 3279-90.

DI GIULIO, M. 2010. Biological evidence against the panspermia theory. *J Theor Biol,* 266**,** 569-72.

DINARELLO, C. A. 2009. Immunological and inflammatory functions of the interleukin-1 family. *Annu Rev Immunol,* 27**,** 519-50.

DOLEI, A. 2006. Endogenous retroviruses and human disease. *Expert Rev Clin Immunol,* 2**,** 149-67.

DONG, L., QI, N., GE, R. M., CAO, C. L., LAN, F. & SHEN, L. 2010. Overexpression of CD133 promotes the phosphorylation of Erk in U87MG human glioblastoma cells. *Neurosci Lett,* 484**,** 210-4.

DONG, X., STOTHARD, P., FORSYTHE, I. J. & WISHART, D. S. 2004. PlasMapper: a web server for drawing and auto-annotating plasmid maps. *Nucleic Acids Res,* 32**,** W660-4.

DRANOFF, G. 2004. Cytokines in cancer pathogenesis and cancer therapy. *Nat Rev Cancer,* 4**,** 11-22.

DREESEN, O. & BRIVANLOU, A. H. 2007. Signaling pathways in cancer and embryonic stem cells. *Stem Cell Rev,* 3**,** 7-17.

DU, L., WANG, H., HE, L., ZHANG, J., NI, B., WANG, X., JIN, H., CAHUZAC, N., MEHRPOUR, M., LU, Y. & CHEN, Q. 2008. CD44 is of functional importance for colorectal cancer stem cells. *Clin Cancer Res,* 14**,** 6751-60.

DUESBERG, P. H., BISTER, K. & VOGT, P. K. 1977. The RNA of avian acute leukemia virus MC29. *Proc Natl Acad Sci U S A,* 74**,** 4320-4.

DUQUE, J., FRESNO, M. & INIGUEZ, M. A. 2005. Expression and function of the nuclear factor of activated T cells in colon carcinoma cells: involvement in the regulation of cyclooxygenase-2. *J Biol Chem,* 280**,** 8686-93.

ECKHOUSE, S., LEWISON, G. & SULLIVAN, R. 2008. Trends in the global funding and activity of cancer research. *Mol Oncol,* 2**,** 20-32.

EICKBUSH, T. H. & FURANO, A. V. 2002. Fruit flies and humans respond differently to retrotransposons. *Curr Opin Genet Dev,* 12**,** 669-74.

ELIYAHU, D., MICHALOVITZ, D., ELIYAHU, S., PINHASI-KIMHI, O. & OREN, M. 1989. Wild-type p53 can inhibit oncogene-mediated focus formation. *Proc Natl Acad Sci U S A,* 86**,** 8763-7.

ENAM, S., DEL VALLE, L., LARA, C., GAN, D. D., ORTIZ-HIDALGO, C., PALAZZO, J. P. & KHALILI, K. 2002. Association of human polyomavirus JCV with colon cancer: evidence for interaction of viral T-antigen and beta-catenin. *Cancer Res,* 62**,** 7093-101.

ERNST, A., AIGNER, M., NAKATA, S., ENGEL, F., SCHLOTTER, M., KLOOR, M., BRAND, K., SCHMITT, S., STEINERT, G., RAHBARI, N., KOCH, M., RADLWIMMER, B., WEITZ, J. & LICHTER, P. 2011. A gene signature distinguishing CD133hi from CD133- colorectal cancer cells: essential role for EGR1 and downstream factors. *Pathology,* 43**,** 220-7.

FAN, H. & JOHNSON, C. 2011. Insertional oncogenesis by non-acute retroviruses: implications for gene therapy. *Viruses,* 3**,** 398-422.

FAN, X., OUYANG, N., TENG, H. & YAO, H. 2011. Isolation and characterization of spheroid cells from the HT29 colon cancer cell line. *Int J Colorectal Dis,* 26**,** 1279-85.

FANG, J. Y. & RICHARDSON, B. C. 2005. The MAPK signalling pathways and colorectal cancer. *Lancet Oncol,* 6**,** 322-7.

FEARON, E. R. 2011. Molecular genetics of colorectal cancer. *Annu Rev Pathol,* 6**,** 479-507.

FEARON, E. R. & VOGELSTEIN, B. 1990. A genetic model for colorectal tumorigenesis. *Cell,* 61**,** 759-67.

FELIX, M. A., ASHE, A., PIFFARETTI, J., WU, G., NUEZ, I., BELICARD, T., JIANG, Y., ZHAO, G., FRANZ, C. J., GOLDSTEIN, L. D., SANROMAN, M., MISKA, E. A. & WANG, D. 2011. Natural and experimental infection of Caenorhabditis nematodes by novel viruses related to nodaviruses. *PLoS Biol,* 9**,** e1000586.

FERLAY, J., SHIN, H. R., BRAY, F., FORMAN, D., MATHERS, C. & PARKIN, D. M. 2010. GLOBOCAN 2008 v2.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 10. Lyon, France: International Agency for Research on Cancer.

FERRANDINA, G., BONANNO, G., PIERELLI, L., PERILLO, A., PROCOLI, A., MARIOTTI, A., CORALLO, M., MARTINELLI, E., RUTELLA, S., PAGLIA, A., ZANNONI, G., MANCUSO, S. & SCAMBIA, G. 2008. Expression of CD133-1 and CD133-2 in ovarian cancer. *Int J Gynecol Cancer,* 18**,** 506-14.

FINLAY, C. A., HINDS, P. W. & LEVINE, A. J. 1989. The p53 proto-oncogene can act as a suppressor of transformation. *Cell,* 57**,** 1083-93.

FINNEGAN, D. J. 1989. Eukaryotic transposable elements and genome evolution. *Trends Genet,* 5**,** 103-7.

FINNEGAN, D. J. 2012. Retrotransposons. *Curr Biol,* 22**,** R432-7.

FOGH, J. T., G. 1975. Human Tumor Cells in Vitro. *In:* FOGH, J. (ed.) *Human Tumor Cells in Vitro.* New York: Plenum Publishing Corp.

FORSMAN, A., YUN, Z., HU, L., UZHAMECKIS, D., JERN, P. & BLOMBERG, J. 2005. Development of broadly targeted human endogenous gammaretroviral pol-based real time PCRs Quantitation of RNA expression in human tissues. *J Virol Methods,* 129**,** 16-30.

FORTERRE, P. 2006. The origin of viruses and their possible roles in major evolutionary transitions. *Virus Res,* 117**,** 5-16.

FORTERRE, P. & PRANGISHVILI, D. 2009. The origin of viruses. *Res Microbiol,* 160**,** 466-72.

FOX, R. 2004. Symbiogenesis. *J R Soc Med,* 97**,** 559.

FRANK, O., GIEHL, M., ZHENG, C., HEHLMANN, R., LEIB-MOSCH, C. & SEIFARTH, W. 2005. Human endogenous retrovirus expression profiles in samples from brains of patients with schizophrenia and bipolar disorders. *J Virol,* 79**,** 10890-901.

FREDERICKS, D. N. & RELMAN, D. A. 1996. Sequence-based identification of microbial pathogens: a reconsideration of Koch's postulates. *Clin Microbiol Rev,* 9**,** 18-33.

FREIMANIS, G., HOOLEY, P., EJTEHADI, H. D., ALI, H. A., VEITCH, A., RYLANCE, P. B., ALAWI, A., AXFORD, J., NEVILL, A., MURRAY, P. G. & NELSON, P. N. 2010. A role for human endogenous retrovirus-K (HML-2) in rheumatoid arthritis: investigating mechanisms of pathogenesis. *Clin Exp Immunol,* 160**,** 340-7.

FUCHS, N. V., LOEWER, S., DALEY, G. Q., IZSVAK, Z., LOWER, J. & LOWER, R. 2013. Human endogenous retrovirus K (HML-2) RNA and protein expression is a marker for human embryonic and induced pluripotent stem cells. *Retrovirology,* 10**,** 115.

FUNG, F. K., CHAN, D. W., LIU, V. W., LEUNG, T. H., CHEUNG, A. N. & NGAN, H. Y. 2012. Increased expression of PITX2 transcription factor contributes to ovarian cancer progression. *PLoS One,* 7**,** e37076.

GALIZIA, G., GEMEI, M., DEL VECCHIO, L., ZAMBOLI, A., DI NOTO, R., MIRABELLI, P., SALVATORE, F., CASTELLANO, P., ORDITURA, M., DE VITA, F., PINTO, M., PIGNATELLI, C. & LIETO, E. 2012. Combined CD133/CD44 expression as a prognostic indicator of disease-free survival in patients with colorectal cancer. *Arch Surg,* 147**,** 18-24.

GALON, J., COSTES, A., SANCHEZ-CABO, F., KIRILOVSKY, A., MLECNIK, B., LAGORCE-PAGES, C., TOSOLINI, M., CAMUS, M., BERGER, A., WIND, P., ZINZINDOHOUE, F., BRUNEVAL, P., CUGNENC, P. H., TRAJANOSKI, Z., FRIDMAN, W. H. & PAGES, F. 2006. Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. *Science,* 313**,** 1960-4.

GARCIA-MONTOJO, M., DOMINGUEZ-MOZO, M., ARIAS-LEAL, A., GARCIA-MARTINEZ, A., DE LAS HERAS, V., CASANOVA, I., FAUCARD, R., GEHIN, N., MADEIRA, A., ARROYO, R., CURTIN, F., ALVAREZ-LAFUENTE, R. & PERRON, H. 2013. The DNA copy number of human endogenous retrovirus-W (MSRV-type) is increased in multiple sclerosis patients and is influenced by gender and disease severity. *PLoS One,* 8**,** e53623.

GARRISON, K. E., JONES, R. B., MEIKLEJOHN, D. A., ANWAR, N., NDHLOVU, L. C., CHAPMAN, J. M., ERICKSON, A. L., AGRAWAL, A., SPOTTS, G., HECHT, F. M., RAKOFF-NAHOUM, S., LENZ, J., OSTROWSKI, M. A. & NIXON, D. F. 2007. T cell responses to human endogenous retroviruses in HIV-1 infection. *PLoS Pathog,* 3**,** e165.

GARSON, J. A., TUKE, P. W., GIRAUD, P., PARANHOS-BACCALA, G. & PERRON, H. 1998. Detection of virion-associated MSRV-RNA in serum of patients with multiple sclerosis. *Lancet,* 351**,** 33.

GASHLER, A. & SUKHATME, V. P. 1995. Early growth response protein 1 (Egr-1): prototype of a zinc-finger family of transcription factors. *Prog Nucleic Acid Res Mol Biol,* 50**,** 191-224.

GAUDIN, P., IJAZ, S., TUKE, P. W., MARCEL, F., PARAZ, A., SEIGNEURIN, J. M., MANDRAND, B., PERRON, H. & GARSON, J. A. 2000. Infrequency of detection of particle-associated MSRV/HERV-W RNA in the synovial fluid of patients with rheumatoid arthritis. *Rheumatology (Oxford),* 39**,** 950-4.

GAVERT, N., SHVAB, A., SHEFFER, M., BEN-SHMUEL, A., HAASE, G., BAKOS, E., DOMANY, E. & BEN-ZE'EV, A. 2013. c-Kit is suppressed in human colon cancer tissue and contributes to L1-mediated metastasis. *Cancer Res,* 73**,** 5754-63.

GERLACH, K., DANIEL, C., LEHR, H. A., NIKOLAEV, A., GERLACH, T., ATREYA, R., ROSE-JOHN, S., NEURATH, M. F. & WEIGMANN, B. 2012. Transcription factor NFATc2 controls the emergence of colon cancer associated with IL-6-dependent colitis. *Cancer Res,* 72**,** 4340-50.

GIAVAZZI, R., DI BERARDINO, C., GAROFALO, A., MOTTA, T., GOBBI, A., SCANZIANI, E., GIUDICI, G., GALLI, M., MAYO, J. G., BARBUI, T. & ET AL. 1995. Establishment of human acute myelogenous leukemia lines secreting interleukin-1 beta in SCID mice. *Int J Cancer,* 61**,** 280-5.

GIFFORD, R. & TRISTEM, M. 2003. The evolution, distribution and diversity of endogenous retroviruses. *Virus Genes,* 26**,** 291-315.

GLINKA, A., WU, W., DELIUS, H., MONAGHAN, A. P., BLUMENSTOCK, C. & NIEHRS, C. 1998. Dickkopf-1 is a member of a new family of secreted proteins and functions in head induction. *Nature,* 391**,** 357-62.

GLOBOCAN 2012. 2012. *GLOBOCAN 2012* [Online]. Available: http://globocan.iarc.fr/.

GOERING, R. V., DOCKRELL, H. M., WAKELIN, D., ZUCKERMAN, M., CHIODINI, P. L., ROITT, I. & MIMS, C. 2008. *Mims' medical microbiology,* Philadelphia, PA Mosby Elsevier.

GOH, H. S., KHINE, K., ELNATAN, J., YAO, J. & SMITH, D. R. 1996. Molecular changes of colorectal cancer in Singapore. *Ann Acad Med Singapore,* 25**,** 3-10.

GREEN, D. R. & KROEMER, G. 2009. Cytoplasmic functions of the tumour suppressor p53. *Nature,* 458**,** 1127-30.

GREGORY, M. A., PHANG, T. L., NEVIANI, P., ALVAREZ-CALDERON, F., EIDE, C. A., O'HARE, T., ZABEREZHNYY, V., WILLIAMS, R. T., DRUKER, B. J., PERROTTI, D. & DEGREGORI, J. 2010. Wnt/Ca2+/NFAT signaling maintains survival of Ph+ leukemia cells upon inhibition of Bcr-Abl. *Cancer Cell,* 18**,** 74-87.

GRETEN, F. R., ECKMANN, L., GRETEN, T. F., PARK, J. M., LI, Z. W., EGAN, L. J., KAGNOFF, M. F. & KARIN, M. 2004. IKKbeta links inflammation and tumorigenesis in a mouse model of colitis-associated cancer. *Cell,* 118**,** 285-96.

GRIVENNIKOV, S. I., GRETEN, F. R. & KARIN, M. 2010. Immunity, inflammation, and cancer. *Cell,* 140**,** 883-99.

GROTEGUT, S., VON SCHWEINITZ, D., CHRISTOFORI, G. & LEHEMBRE, F. 2006. Hepatocyte growth factor induces cell scattering through MAPK/Egr-1-mediated upregulation of Snail. *EMBO J,* 25**,** 3534-45.

GRYFE, R. 2009. Inherited colorectal cancer syndromes. *Clin Colon Rectal Surg,* 22**,** 198-208.

GULIYEV, M., YILMAZ, S., SAHIN, K., MARAKLI, S. & GOZUKIRMIZI, N. 2013. Human endogenous retrovirus-H insertion screening. *Mol Med Rep,* 7**,** 1305-9.

HAIG, D. 2012. Retroviruses and the placenta. *Curr Biol,* 22**,** R609-13.

HAMPTON, T. 2006. New markers may help predict prostate cancer relapse risk. *JAMA,* 295**,** 2234-8.

HAN, J., JIANG, Y., LI, Z., KRAVCHENKO, V. V. & ULEVITCH, R. J. 1997. Activation of the transcription factor MEF2C by the MAP kinase p38 in inflammation. *Nature,* 386**,** 296-9.

HANAHAN, D. & WEINBERG, R. A. 2000. The hallmarks of cancer. *Cell,* 100**,** 57-70.

HANAHAN, D. & WEINBERG, R. A. 2011. Hallmarks of cancer: the next generation. *Cell,* 144**,** 646-74.

HANKE, K., CHUDAK, C., KURTH, R. & BANNERT, N. 2013. The Rec protein of HERV-K(HML-2) upregulates androgen receptor activity by binding to the human small glutamine-rich tetratricopeptide repeat protein (hSGT). *Int J Cancer,* 132**,** 556-67.

HARADA, F., TSUKADA, N. & KATO, N. 1987. Isolation of three kinds of human endogenous retrovirus-like sequences using tRNA(Pro) as a probe. *Nucleic Acids Res,* 15**,** 9153-62.

HARADA, N., MIZOI, T., KINOUCHI, M., HOSHI, K., ISHII, S., SHIIBA, K., SASAKI, I. & MATSUNO, S. 2001. Introduction of antisense CD44S CDNA down-regulates expression of overall CD44 isoforms and inhibits tumor growth and metastasis in highly metastatic colon carcinoma cells. *Int J Cancer,* 91**,** 67-75.

HARADA, S., SMITH, R. M., SMITH, J. A., WHITE, M. F. & JARETT, L. 1996. Insulin-induced egr-1 and c-fos expression in 32D cells requires insulin receptor, Shc, and mitogen-activated protein kinase, but not insulin receptor substrate-1 and phosphatidylinositol 3-kinase activation. *J Biol Chem,* 271**,** 30222-6.

HARRIS, S. L. & LEVINE, A. J. 2005. The p53 pathway: positive and negative feedback loops. *Oncogene,* 24**,** 2899-908.

HATZIVASSILIOU, E., MILLER, W. E., RAAB-TRAUB, N., KIEFF, E. & MOSIALOS, G. 1998. A fusion of the EBV latent membrane protein-1 (LMP1) transmembrane domains to the CD40 cytoplasmic domain is similar to LMP1 in constitutive activation of epidermal growth factor receptor expression, nuclear factor-kappa B, and stress-activated protein kinase. *J Immunol,* 160**,** 1116-21.

HAVERKOS, H. W. 2004. Viruses, chemicals and co-carcinogenesis. *Oncogene,* 23**,** 6492-9.

HAYWARD, S. D., LIU, J. & FUJIMURO, M. 2006. Notch and Wnt signaling: mimicry and manipulation by gamma herpesviruses. *Sci STKE,* 2006**,** re4.

HAYWARD, W. S., NEEL, B. G. & ASTRIN, S. M. 1981. Activation of a cellular onc gene by promoter insertion in ALV-induced lymphoid leukosis. *Nature,* 290**,** 475-80.

HE, X. X., CHEN, K., YANG, J., LI, X. Y., GAN, H. Y., LIU, C. Y., COLEMAN, T. R. & AL-ABED, Y. 2009. Macrophage migration inhibitory factor promotes colorectal cancer. *Mol Med,* 15**,** 1-10.

HERBST, H., SAUTER, M. & MUELLER-LANTZSCH, N. 1996. Expression of human endogenous retrovirus K elements in germ cell and trophoblastic tumors. *Am J Pathol,* 149**,** 1727-35.

HERISHANU, Y., GIBELLINI, F., NJUGUNA, N., HAZAN-HALEVY, I., FAROOQUI, M., BERN, S., KEYVANFAR, K., LEE, E., WILSON, W. & WIESTNER, A. 2011. Activation of CD44, a receptor for extracellular matrix components, protects

chronic lymphocytic leukemia cells from spontaneous and drug induced apoptosis through MCL-1. *Leuk Lymphoma,* 52**,** 1758-69.

HERMANN, P. C., HUBER, S. L., HERRLER, T., AICHER, A., ELLWART, J. W., GUBA, M., BRUNS, C. J. & HEESCHEN, C. 2007. Distinct populations of cancer stem cells determine tumor growth and metastatic activity in human pancreatic cancer. *Cell Stem Cell,* 1**,** 313-23.

HILL, A. B. 1965. THE ENVIRONMENT AND DISEASE: ASSOCIATION OR CAUSATION? *Proc R Soc Med,* 58**,** 295-300.

HIPSKIND, R. A., BUSCHER, D., NORDHEIM, A. & BACCARINI, M. 1994. Ras/MAP kinase-dependent and -independent signaling pathways target distinct ternary complex factors. *Genes Dev,* 8**,** 1803-16.

HIROSE, H., ISHII, H., MIMORI, K., TANAKA, F., TAKEMASA, I., MIZUSHIMA, T., IKEDA, M., YAMAMOTO, H., SEKIMOTO, M., DOKI, Y. & MORI, M. 2011. The significance of PITX2 overexpression in human colorectal cancer. *Ann Surg Oncol,* 18**,** 3005-12.

HIRSCH, D., BARKER, N., MCNEIL, N., HU, Y., CAMPS, J., MCKINNON, K., CLEVERS, H., RIED, T. & GAISER, T. 2013. LGR5 positivity defines stem-like cells in colorectal cancer. *Carcinogenesis*.

HISHIKAWA, T., OGASAWARA, H., KANEKO, H., SHIRASAWA, T., MATSUURA, Y., SEKIGAWA, I., TAKASAKI, Y., HASHIMOTO, H., HIROSE, S., HANDA, S., NAGASAWA, R. & MARUYAMA, N. 1997. Detection of antibodies to a recombinant gag protein derived from human endogenous retrovirus clone 4-1 in autoimmune diseases. *Viral Immunol,* 10**,** 137-47.

HODGE, C., LIAO, J., STOFEGA, M., GUAN, K., CARTER-SU, C. & SCHWARTZ, J. 1998. Growth hormone stimulates phosphorylation and activation of elk-1 and expression of c-fos, egr-1, and junB through activation of extracellular signal-regulated kinases 1 and 2. *J Biol Chem,* 273**,** 31327-36.

HOHN, O., HANKE, K. & BANNERT, N. 2013. HERV-K(HML-2), the Best Preserved Family of HERVs: Endogenization, Expression, and Implications in Health and Disease. *Front Oncol,* 3**,** 246.

HOMMINGA, I., PIETERS, R., LANGERAK, A. W., DE ROOI, J. J., STUBBS, A., VERSTEGEN, M., VUERHARD, M., BUIJS-GLADDINES, J., KOOI, C., KLOUS, P., VAN VLIERBERGHE, P., FERRANDO, A. A., CAYUELA, J. M., VERHAAF, B., BEVERLOO, H. B., HORSTMANN, M., DE HAAS, V., WIEKMEIJER, A. S., PIKE-OVERZET, K., STAAL, F. J., DE LAAT, W., SOULIER, J., SIGAUX, F. & MEIJERINK, J. P. 2011. Integrated transcript and genome analyses reveal NKX2-1 and MEF2C as potential oncogenes in T cell acute lymphoblastic leukemia. *Cancer Cell,* 19**,** 484-97.

HON, G. M., ERASMUS, R. T. & MATSHA, T. 2013. Multiple sclerosis-associated retrovirus and related human endogenous retrovirus-W in patients with multiple sclerosis: A literature review. *J Neuroimmunol,* 263**,** 8-12.

HONG, Y., HO, K. S., EU, K. W. & CHEAH, P. Y. 2007. A susceptibility gene set for early onset colorectal cancer that integrates diverse signaling pathways: implication for tumorigenesis. *Clin Cancer Res,* 13**,** 1107-14.

HORN, H. F. & VOUSDEN, K. H. 2007. Coping with stress: multiple ways to activate p53. *Oncogene,* 26**,** 1306-16.

HORSLEY, V. & PAVLATH, G. K. 2002. NFAT: ubiquitous regulator of cell differentiation and adaptation. *J Cell Biol,* 156**,** 771-4.

HORST, D., KRIEGL, L., ENGEL, J., JUNG, A. & KIRCHNER, T. 2009a. CD133 and nuclear beta-catenin: the marker combination to detect high risk cases of low stage colorectal cancer. *Eur J Cancer,* 45**,** 2034-40.

HORST, D., KRIEGL, L., ENGEL, J., KIRCHNER, T. & JUNG, A. 2009b. Prognostic significance of the cancer stem cell markers CD133, CD44, and CD166 in colorectal cancer. *Cancer Invest,* 27**,** 844-50.

HORST, D., SCHEEL, S. K., LIEBMANN, S., NEUMANN, J., MAATZ, S., KIRCHNER, T. & JUNG, A. 2009c. The cancer stem cell marker CD133 has high prognostic impact but unknown functional relevance for the metastasis of human colon cancer. *J Pathol,* 219**,** 427-34.

HOSHINO, R., TANIMURA, S., WATANABE, K., KATAOKA, T. & KOHNO, M. 2001. Blockade of the extracellular signal-regulated kinase pathway induces marked G1 cell cycle arrest and apoptosis in tumor cells in which the pathway is constitutively activated: up-regulation of p27(Kip1). *J Biol Chem,* 276**,** 2686-92.

HU, L., HORNUNG, D., KUREK, R., OSTMAN, H., BLOMBERG, J. & BERGQVIST, A. 2006. Expression of human endogenous gammaretroviral sequences in endometriosis and ovarian cancer. *AIDS Res Hum Retroviruses,* 22**,** 551-7.

HU, M. C., QIU, W. R., WANG, X., MEYER, C. F. & TAN, T. H. 1996. Human HPK1, a novel human hematopoietic progenitor kinase that activates the JNK/SAPK kinase cascade. *Genes Dev,* 10**,** 2251-64.

HUANG, R. P., FAN, Y., DEBELLE, I., NI, Z., MATHENY, W. & ADAMSON, E. D. 1998a. Egr-1 inhibits apoptosis during the UV response: correlation of cell survival with Egr-1 phosphorylation. *Cell Death Differ,* 5**,** 96-106.

HUANG, S., JEAN, D., LUCA, M., TAINSKY, M. A. & BAR-ELI, M. 1998b. Loss of AP-2 results in downregulation of c-KIT and enhancement of melanoma tumorigenicity and metastasis. *EMBO J,* 17**,** 4358-69.

HUANG, W. J., LIU, Z. C., WEI, W., WANG, G. H., WU, J. G. & ZHU, F. 2006. Human endogenous retroviral pol RNA and protein detected and identified in the blood of individuals with schizophrenia. *Schizophr Res,* 83**,** 193-9.

HWANG, W. L., YANG, M. H., TSAI, M. L., LAN, H. Y., SU, S. H., CHANG, S. C., TENG, H. W., YANG, S. H., LAN, Y. T., CHIOU, S. H. & WANG, H. W. 2011. SNAIL regulates interleukin-8 expression, stem cell-like activity, and tumorigenicity of human colorectal carcinoma cells. *Gastroenterology,* 141**,** 279-91, 291 e1-5.

IETA, K., TANAKA, F., HARAGUCHI, N., KITA, Y., SAKASHITA, H., MIMORI, K., MATSUMOTO, T., INOUE, H., KUWANO, H. & MORI, M. 2008. Biological and genetic characteristics of tumor-initiating cells in colon cancer. *Ann Surg Oncol,* 15**,** 638-48.

INTERNATIONAL COMMITTEE ON TAXONOMY OF VIRUSES. 2012. *The official ICTV 2011 taxonomy* [Online]. International Committee on Taxonomy of Viruses,. Available: http://ictvonline.org/index.asp [Accessed 30 December 2012 2012].

ITZKOWITZ, S. H. & YIO, X. 2004. Inflammation and cancer IV. Colorectal cancer in inflammatory bowel disease: the role of inflammation. *Am J Physiol Gastrointest Liver Physiol,* 287**,** G7-17.

JAECKEL, E., MANNS, M. & VON HERRATH, M. 2002. Viruses and diabetes. *Ann N Y Acad Sci,* 958**,** 7-25.

JAULIAC, S., LOPEZ-RODRIGUEZ, C., SHAW, L. M., BROWN, L. F., RAO, A. & TOKER, A. 2002. The role of NFAT transcription factors in integrin-mediated carcinoma invasion. *Nat Cell Biol,* 4**,** 540-4.

JHA, A. R., PILLAI, S. K., YORK, V. A., SHARP, E. R., STORM, E. C., WACHTER, D. J., MARTIN, J. N., DEEKS, S. G., ROSENBERG, M. G., NIXON, D. F. & GARRISON, K. E. 2009. Cross-sectional dating of novel haplotypes of HERV-K 113 and

HERV-K 115 indicate these proviruses originated in Africa before Homo sapiens. *Mol Biol Evol,* 26**,** 2617-26.

JIN, S. & LEVINE, A. J. 2001. The p53 functional circuit. *J Cell Sci,* 114**,** 4139-40.

JOHANSSON, M., DENARDO, D. G. & COUSSENS, L. M. 2008. Polarized immune responses differentially regulate cancer development. *Immunol Rev,* 222**,** 145-54.

KAISER, S., PARK, Y. K., FRANKLIN, J. L., HALBERG, R. B., YU, M., JESSEN, W. J., FREUDENBERG, J., CHEN, X., HAIGIS, K., JEGGA, A. G., KONG, S., SAKTHIVEL, B., XU, H., REICHLING, T., AZHAR, M., BOIVIN, G. P., ROBERTS, R. B., BISSAHOYO, A. C., GONZALES, F., BLOOM, G. C., ESCHRICH, S., CARTER, S. L., ARONOW, J. E., KLEIMEYER, J., KLEIMEYER, M., RAMASWAMY, V., SETTLE, S. H., BOONE, B., LEVY, S., GRAFF, J. M., DOETSCHMAN, T., GRODEN, J., DOVE, W. F., THREADGILL, D. W., YEATMAN, T. J., COFFEY, R. J., JR. & ARONOW, B. J. 2007. Transcriptional recapitulation and subversion of embryonic colon development by mouse colon tumor models and human colon cancer. *Genome Biol,* 8**,** R131.

KALLAND, K. H., KE, X. S. & OYAN, A. M. 2009. Tumour virology--history, status and future challenges. *APMIS,* 117**,** 382-99.

KAMP, D. W., SHACTER, E. & WEITZMAN, S. A. 2011. Chronic inflammation and cancer: the role of the mitochondria. *Oncology (Williston Park),* 25**,** 400-10, 413.

KANEHISA, M. 2013. Molecular network analysis of diseases and drugs in KEGG. *Methods Mol Biol,* 939**,** 263-75.

KANEHISA, M., GOTO, S., SATO, Y., FURUMICHI, M. & TANABE, M. 2012. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res,* 40**,** D109-14.

KANWAR, S. S., YU, Y., NAUTIYAL, J., PATEL, B. B. & MAJUMDAR, A. P. 2010. The Wnt/beta-catenin pathway regulates growth and maintenance of colonospheres. *Mol Cancer,* 9**,** 212.

KARLSSON, H., BACHMANN, S., SCHRODER, J., MCARTHUR, J., TORREY, E. F. & YOLKEN, R. H. 2001. Retroviral RNA identified in the cerebrospinal fluids and brains of individuals with schizophrenia. *Proc Natl Acad Sci U S A,* 98**,** 4634-9.

KARLSSON, H., SCHRODER, J., BACHMANN, S., BOTTMER, C. & YOLKEN, R. H. 2004. HERV-W-related RNA detected in plasma from individuals with recent-onset schizophrenia or schizoaffective disorder. *Mol Psychiatry,* 9**,** 12-3.

KATOH, I. & KURATA, S. I. 2013. Association of Endogenous Retroviruses and Long Terminal Repeats with Human Disorders. *Front Oncol,* 3**,** 234.

KATOH, Y. & KATOH, M. 2007. Comparative genomics on PROM1 gene encoding stem cell marker CD133. *Int J Mol Med,* 19**,** 967-70.

KHEIRELSEID, E., MILLER, N. & KERIN, M. 2013. Molecular biology of colorectal cancer: Review of the literature. *American Journal of Molecular Biology,* 3**,** 72-80.

KHIEM, D., CYSTER, J. G., SCHWARZ, J. J. & BLACK, B. L. 2008. A p38 MAPK-MEF2C pathway regulates B-cell proliferation. *Proc Natl Acad Sci U S A,* 105**,** 17067-72.

KIEFER, F., TIBBLES, L. A., ANAFI, M., JANSSEN, A., ZANKE, B. W., LASSAM, N., PAWSON, T., WOODGETT, J. R. & ISCOVE, N. N. 1996. HPK1, a hematopoietic protein kinase activating the SAPK/JNK pathway. *EMBO J,* 15**,** 7013-25.

KIESER, A., KILGER, E., GIRES, O., UEFFING, M., KOLCH, W. & HAMMERSCHMIDT, W. 1997. Epstein-Barr virus latent membrane protein-1 triggers AP-1 activity via the c-Jun N-terminal kinase cascade. *EMBO J,* 16**,** 6478-85.

KIKUCHI, A., YAMAMOTO, H., SATO, A. & MATSUMOTO, S. 2011. New insights into the mechanism of Wnt signaling pathway activation. *Int Rev Cell Mol Biol,* 291**,** 21-71.

KIM, K. A., KAKITANI, M., ZHAO, J., OSHIMA, T., TANG, T., BINNERTS, M., LIU, Y., BOYLE, B., PARK, E., EMTAGE, P., FUNK, W. D. & TOMIZUKA, K. 2005. Mitogenic influence of human R-spondin1 on the intestinal epithelium. *Science,* 309**,** 1256-9.

KIM, Y. S., HWAN, J. D., BAE, S., BAE, D. H. & SHICK, W. A. 2010. Identification of differentially expressed genes using an annealing control primer system in stage III serous ovarian carcinoma. *BMC Cancer,* 10**,** 576.

KING, A. M. Q., LEFKOWITZ, E., ADAMS, M. J. & CARSTENS, E. B. 2011. *Virus Taxonomy. Ninth Report of the International Committee on Taxonomy of Viruses*, Academic Press, Elsevier.

KIOUSSI, C., BRIATA, P., BAEK, S. H., ROSE, D. W., HAMBLET, N. S., HERMAN, T., OHGI, K. A., LIN, C., GLEIBERMAN, A., WANG, J., BRAULT, V., RUIZ-LOZANO, P., NGUYEN, H. D., KEMLER, R., GLASS, C. K., WYNSHAW-BORIS, A. & ROSENFELD, M. G. 2002. Identification of a Wnt/Dvl/beta-Catenin --> Pitx2 pathway mediating cell-type-specific proliferation during development. *Cell,* 111**,** 673-85.

KLEFFEL, S. & SCHATTON, T. 2013. Tumor dormancy and cancer stem cells: two sides of the same coin? *Adv Exp Med Biol,* 734**,** 145-79.

KOBAYASHI, M., KIM, J. Y., CAMARENA, V., ROEHM, P. C., CHAO, M. V., WILSON, A. C. & MOHR, I. 2012. A primary neuron culture system for the study of herpes simplex virus latency and reactivation. *J Vis Exp*.

KOEHNE, C. H. & DUBOIS, R. N. 2004. COX-2 inhibition and colorectal cancer. *Semin Oncol,* 31**,** 12-21.

KOHN, A. D. & MOON, R. T. 2005. Wnt and calcium signaling: beta-catenin-independent pathways. *Cell Calcium,* 38**,** 439-46.

KOMURIAN-PRADEL, F., PARANHOS-BACCALA, G., BEDIN, F., OUNANIAN-PARAZ, A., SODOYER, M., OTT, C., RAJOHARISON, A., GARCIA, E., MALLET, F., MANDRAND, B. & PERRON, H. 1999. Molecular cloning and characterization of MSRV-related sequences associated with retrovirus-like particles. *Virology,* 260**,** 1-9.

KRAUSS, G. 2008. *Biochemistry of signal transduction and regulation,* Weinheim, WILEY-VCH Verlag GmbH & Co.

KRAUTWALD, S. 1998. IL-16 activates the SAPK signaling pathway in CD4+ macrophages. *J Immunol,* 160**,** 5874-9.

KRISHNA, M. & NARANG, H. 2008. The complexity of mitogen-activated protein kinases (MAPKs) made simple. *Cell Mol Life Sci,* 65**,** 3525-44.

KRONES-HERZIG, A., MITTAL, S., YULE, K., LIANG, H., ENGLISH, C., URCIS, R., SONI, T., ADAMSON, E. D. & MERCOLA, D. 2005. Early growth response 1 acts as a tumor suppressor in vivo and in vitro via regulation of p53. *Cancer Res,* 65**,** 5133-43.

KUMAR, V., ABBAS, A. K., FAUSTO, N. & ASTER, J. 2009. Acute and Chronic Inflammation. *Robbins and Cotran Pathologic Basis of Disease.* 8 ed. Philadelphia, PA: SAUNDERS ELSEVIER.

LA MANTIA, G., MAGLIONE, D., PENGUE, G., DI CRISTOFANO, A., SIMEONE, A., LANFRANCONE, L. & LANIA, L. 1991. Identification and characterization of

novel human endogenous retroviral sequences prefentially expressed in undifferentiated embryonal carcinoma cells. *Nucleic Acids Res,* 19**,** 1513-20.

LAI, M. M., HU, S. S. & VOGT, P. K. 1979. Avian erythroblastosis virus: transformation-specific sequences form a contiguous segment of 3.25 kb located in the middle of the 6-kb genome. *Virology,* 97**,** 366-77.

LAMPRECHT, B., BONIFER, C. & MATHAS, S. 2010a. Repeat-element driven activation of proto-oncogenes in human malignancies. *Cell Cycle,* 9**,** 4276-81.

LAMPRECHT, B., WALTER, K., KREHER, S., KUMAR, R., HUMMEL, M., LENZE, D., KOCHERT, K., BOUHLEL, M. A., RICHTER, J., SOLER, E., STADHOUDERS, R., JOHRENS, K., WURSTER, K. D., CALLEN, D. F., HARTE, M. F., GIEFING, M., BARLOW, R., STEIN, H., ANAGNOSTOPOULOS, I., JANZ, M., COCKERILL, P. N., SIEBERT, R., DORKEN, B., BONIFER, C. & MATHAS, S. 2010b. Derepression of an endogenous long terminal repeat activates the CSF1R proto-oncogene in human lymphoma. *Nat Med,* 16**,** 571-9, 1p following 579.

LAMPUGNANI, M. G. 1999. Cell migration into a wounded area in vitro. *Methods Mol Biol,* 96**,** 177-82.

LANDER, E. S., LINTON, L. M., BIRREN, B., NUSBAUM, C., ZODY, M. C., BALDWIN, J., DEVON, K., DEWAR, K., DOYLE, M., FITZHUGH, W., FUNKE, R., GAGE, D., HARRIS, K., HEAFORD, A., HOWLAND, J., KANN, L., LEHOCZKY, J., LEVINE, R., MCEWAN, P., MCKERNAN, K., MELDRIM, J., MESIROV, J. P., MIRANDA, C., MORRIS, W., NAYLOR, J., RAYMOND, C., ROSETTI, M., SANTOS, R., SHERIDAN, A., SOUGNEZ, C., STANGE-THOMANN, N., STOJANOVIC, N., SUBRAMANIAN, A., WYMAN, D., ROGERS, J., SULSTON, J., AINSCOUGH, R., BECK, S., BENTLEY, D., BURTON, J., CLEE, C., CARTER, N., COULSON, A., DEADMAN, R., DELOUKAS, P., DUNHAM, A., DUNHAM, I., DURBIN, R., FRENCH, L., GRAFHAM, D., GREGORY, S., HUBBARD, T., HUMPHRAY, S., HUNT, A., JONES, M., LLOYD, C., MCMURRAY, A., MATTHEWS, L., MERCER, S., MILNE, S., MULLIKIN, J. C., MUNGALL, A., PLUMB, R., ROSS, M., SHOWNKEEN, R., SIMS, S., WATERSTON, R. H., WILSON, R. K., HILLIER, L. W., MCPHERSON, J. D., MARRA, M. A., MARDIS, E. R., FULTON, L. A., CHINWALLA, A. T., PEPIN, K. H., GISH, W. R., CHISSOE, S. L., WENDL, M. C., DELEHAUNTY, K. D., MINER, T. L., DELEHAUNTY, A., KRAMER, J. B., COOK, L. L., FULTON, R. S., JOHNSON, D. L., MINX, P. J., CLIFTON, S. W., HAWKINS, T., BRANSCOMB, E., PREDKI, P., RICHARDSON, P., WENNING, S., SLEZAK, T., DOGGETT, N., CHENG, J. F., OLSEN, A., LUCAS, S., ELKIN, C., UBERBACHER, E., FRAZIER, M., et al. 2001. Initial sequencing and analysis of the human genome. *Nature,* 409**,** 860-921.

LANE, D. P. 1992. Cancer. p53, guardian of the genome. *Nature,* 358**,** 15-6.

LARSSON, E., KATO, N. & COHEN, M. 1989. Human endogenous proviruses. *Curr Top Microbiol Immunol,* 148**,** 115-32.

LASCORZ, J., FORSTI, A., CHEN, B., BUCH, S., STEINKE, V., RAHNER, N., HOLINSKI-FEDER, E., MORAK, M., SCHACKERT, H. K., GORGENS, H., SCHULMANN, K., GOECKE, T., KLOOR, M., ENGEL, C., BUTTNER, R., KUNKEL, N., WEIRES, M., HOFFMEISTER, M., PARDINI, B., NACCARATI, A., VODICKOVA, L., NOVOTNY, J., SCHREIBER, S., KRAWCZAK, M., BRORING, C. D., VOLZKE, H., SCHAFMAYER, C., VODICKA, P., CHANG-CLAUDE, J., BRENNER, H., BURWINKEL, B., PROPPING, P., HAMPE, J. & HEMMINKI, K. 2010. Genome-wide association study for colorectal cancer identifies risk polymorphisms in German familial cases and implicates MAPK signalling pathways in disease susceptibility. *Carcinogenesis,* 31**,** 1612-9.

LEBOYER, M., TAMOUZA, R., CHARRON, D., FAUCARD, R. & PERRON, H. 2011. Human endogenous retrovirus type W (HERV-W) in schizophrenia: A new avenue of research at the gene-environment interface. *World J Biol Psychiatry*.

LEBOYER, M., TAMOUZA, R., CHARRON, D., FAUCARD, R. & PERRON, H. 2013. Human endogenous retrovirus type W (HERV-W) in schizophrenia: a new avenue of research at the gene-environment interface. *World J Biol Psychiatry,* 14**,** 80-90.

LEUNG, D. C. & LORINCZ, M. C. 2012. Silencing of endogenous retroviruses: when and why do histone marks predominate? *Trends Biochem Sci,* 37**,** 127-33.

LEVINE, A. J. 1997. p53, the cellular gatekeeper for growth and division. *Cell,* 88**,** 323-31.

LEVINE, A. J., FINLAY, C. A. & HINDS, P. W. 2004. P53 is a tumor suppressor gene. *Cell,* 116**,** S67-9, 1 p following S69.

LI, L., DRAGULEV, B., ZIGRINO, P., MAUCH, C. & FOX, J. W. 2009. The invasive potential of human melanoma cell lines correlates with their ability to alter fibroblast gene expression in vitro and the stromal microenvironment in vivo. *Int J Cancer,* 125**,** 1796-804.

LI, X., LEWIS, M. T., HUANG, J., GUTIERREZ, C., OSBORNE, C. K., WU, M. F., HILSENBECK, S. G., PAVLICK, A., ZHANG, X., CHAMNESS, G. C., WONG, H., ROSEN, J. & CHANG, J. C. 2008. Intrinsic resistance of tumorigenic breast cancer cells to chemotherapy. *J Natl Cancer Inst,* 100**,** 672-9.

LIANG, C. C., PARK, A. Y. & GUAN, J. L. 2007a. In vitro scratch assay: a convenient and inexpensive method for analysis of cell migration in vitro. *Nat Protoc,* 2**,** 329-33.

LIANG, Q., DING, J., XU, R., XU, Z. & ZHENG, S. 2009. Identification of a novel human endogenous retrovirus and promoter activity of its 5' U3. *Biochem Biophys Res Commun,* 382**,** 468-72.

LIANG, Q. Y., XU, Z. F., XU, R. Z., ZHENG, S. & DING, J. Y. 2007b. [Deletion of the env region in HERV-H-X gene and its expression in colon cancer]. *Ai Zheng,* 26**,** 952-6.

LIEBRICH, M., GUO, L. H., SCHLUESENER, H. J., SCHWAB, J. M., DIETZ, K., WILL, B. E. & MEYERMANN, R. 2007. Expression of interleukin-16 by tumor-associated macrophages/activated microglia in high-grade astrocytic brain tumors. *Arch Immunol Ther Exp (Warsz),* 55**,** 41-7.

LIM, C. P., JAIN, N. & CAO, X. 1998. Stress-induced immediate-early gene, egr-1, involves activation of p38/JNK1. *Oncogene,* 16**,** 2915-26.

LIN, P. Y., FUNG, C. Y., CHANG, F. P., HUANG, W. S., CHEN, W. C., WANG, J. Y. & CHANG, D. 2008. Prevalence and genotype identification of human JC virus in colon cancer in Taiwan. *J Med Virol,* 80**,** 1828-34.

LIPKIN, M., REDDY, B., NEWMARK, H. & LAMPRECHT, S. A. 1999. Dietary factors in human colorectal cancer. *Annu Rev Nutr,* 19**,** 545-86.

LIU, C., RANGNEKAR, V. M., ADAMSON, E. & MERCOLA, D. 1998. Suppression of growth and transformation and induction of apoptosis by EGR-1. *Cancer Gene Ther,* 5**,** 3-28.

LIU, Y. & BODMER, W. F. 2006. Analysis of P53 mutations and their expression in 56 colorectal cancer cell lines. *Proc Natl Acad Sci U S A,* 103**,** 976-81.

LIVAK, K. J. & SCHMITTGEN, T. D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods,* 25**,** 402-8.

LO, L. W., CHENG, J. J., CHIU, J. J., WUNG, B. S., LIU, Y. C. & WANG, D. L. 2001. Endothelial exposure to hypoxia induces Egr-1 expression involving PKCalpha-mediated Ras/Raf-1/ERK1/2 pathway. *J Cell Physiol,* 188**,** 304-12.

LOBO, N. A., SHIMONO, Y., QIAN, D. & CLARKE, M. F. 2007. The biology of cancer stem cells. *Annu Rev Cell Dev Biol,* 23**,** 675-99.

LOGAN, C. Y. & NUSSE, R. 2004. The Wnt signaling pathway in development and disease. *Annu Rev Cell Dev Biol,* 20**,** 781-810.

LONGO, D. L. F. A. S. 2010. The Human Retroviruses. *In:* KASPER, D. & FAUCI, A. (eds.) *Harrison's Infectious Diseases.* 1 ed. New York: McGraw-Hill.

LOWER, R., BOLLER, K., HASENMAIER, B., KORBMACHER, C., MULLER-LANTZSCH, N., LOWER, J. & KURTH, R. 1993. Identification of human endogenous retroviruses with complex mRNA expression and particle formation. *Proc Natl Acad Sci U S A,* 90**,** 4480-4.

LOWER, R., LOWER, J. & KURTH, R. 1996. The viruses in all of us: characteristics and biological significance of human endogenous retrovirus sequences. *Proc Natl Acad Sci U S A,* 93**,** 5177-84.

LU, H., OUYANG, W. & HUANG, C. 2006. Inflammation, a key event in cancer development. *Mol Cancer Res,* 4**,** 221-33.

LUST, J. A. & DONOVAN, K. A. 1999. The role of interleukin-1 beta in the pathogenesis of multiple myeloma. *Hematol Oncol Clin North Am,* 13**,** 1117-25.

MA, W., XIA, C., LING, P., QIU, M., LUO, Y., TAN, T. H. & LIU, M. 2001. Leukocyte-specific adaptor protein Grap2 interacts with hematopoietic progenitor kinase 1 (HPK1) to activate JNK signaling pathway in T lymphocytes. *Oncogene,* 20**,** 1703-14.

MACFARLAN, T. S., GIFFORD, W. D., DRISCOLL, S., LETTIERI, K., ROWE, H. M., BONANOMI, D., FIRTH, A., SINGER, O., TRONO, D. & PFAFF, S. L. 2012. Embryonic stem cell potency fluctuates with endogenous retrovirus activity. *Nature,* 487**,** 57-63.

MAFFINI, M. V., SOTO, A. M., SONNENSCHEIN, C., PAPADOPOULOS, N. & THEOHARIDES, T. C. 2008. Lack of c-kit receptor promotes mammary tumors in N-nitrosomethylurea-treated Ws/Ws rats. *Cancer Cell Int,* 8**,** 5.

MAGER, D. L. & FREEMAN, J. D. 1987. Human endogenous retroviruslike genome with type C pol sequences and gag sequences related to human T-cell lymphotropic viruses. *J Virol,* 61**,** 4060-6.

MAGER, D. L. & FREEMAN, J. D. 1995. HERV-H endogenous retroviruses: presence in the New World branch but amplification in the Old World primate lineage. *Virology,* 213**,** 395-404.

MAGER, D. L. & HENTHORN, P. S. 1984. Identification of a retrovirus-like repetitive element in human DNA. *Proc Natl Acad Sci U S A,* 81**,** 7510-4.

MAHALINGAM, D., NATONI, A., KEANE, M., SAMALI, A. & SZEGEZDI, E. 2010. Early growth response-1 is a regulator of DR5-induced apoptosis in colon cancer cells. *Br J Cancer,* 102**,** 754-64.

MAIER, S., NIMMRICH, I., KOENIG, T., EPPENBERGER-CASTORI, S., BOHLMANN, I., PARADISO, A., SPYRATOS, F., THOMSSEN, C., MUELLER, V., NAHRIG, J., SCHITTULLI, F., KATES, R., LESCHE, R., SCHWOPE, I., KLUTH, A., MARX, A., MARTENS, J. W., FOEKENS, J. A., SCHMITT, M. & HARBECK, N. 2007. DNA-methylation of the homeodomain transcription factor PITX2 reliably predicts risk of distant disease recurrence in tamoxifen-treated, node-negative breast cancer patients--Technical and clinical validation in a multi-centre

setting in collaboration with the European Organisation for Research and Treatment of Cancer (EORTC) PathoBiology group. *Eur J Cancer,* 43**,** 1679-86.

MAKSAKOVA, I. A., MAGER, D. L. & REISS, D. 2008. Keeping active endogenous retroviral-like elements in check: the epigenetic perspective. *Cell Mol Life Sci,* 65**,** 3329-47.

MAKSAKOVA, I. A., ROMANISH, M. T., GAGNIER, L., DUNN, C. A., VAN DE LAGEMAAT, L. N. & MAGER, D. L. 2006. Retroviral elements and their hosts: insertional mutagenesis in the mouse germ line. *PLoS Genet,* 2**,** e2.

MANGENEY, M., RENARD, M., SCHLECHT-LOUF, G., BOUALLAGA, I., HEIDMANN, O., LETZELTER, C., RICHAUD, A., DUCOS, B. & HEIDMANN, T. 2007. Placental syncytins: Genetic disjunction between the fusogenic and immunosuppressive activity of retroviral envelope proteins. *Proc Natl Acad Sci U S A,* 104**,** 20534-9.

MANGHERA, M. & DOUVILLE, R. N. 2013. Endogenous retrovirus-K promoter: a landing strip for inflammatory transcription factors? *Retrovirology,* 10**,** 16.

MANTOVANI, A., ALLAVENA, P., SICA, A. & BALKWILL, F. 2008. Cancer-related inflammation. *Nature,* 454**,** 436-44.

MAO, B., WU, W., LI, Y., HOPPE, D., STANNEK, P., GLINKA, A. & NIEHRS, C. 2001. LDL-receptor-related protein 6 is a receptor for Dickkopf proteins. *Nature,* 411**,** 321-5.

MARCHI, E., KANAPIN, A., BYOTT, M., MAGIORKINIS, G. & BELSHAW, R. 2013. Neanderthal and Denisovan retroviruses in modern humans. *Curr Biol,* 23**,** R994-5.

MARGUERAT, S., WANG, W. Y., TODD, J. A. & CONRAD, B. 2004. Association of human endogenous retrovirus K-18 polymorphisms with type 1 diabetes. *Diabetes,* 53**,** 852-4.

MARIOTTO, A. B., YABROFF, K. R., SHAO, Y., FEUER, E. J. & BROWN, M. L. 2011. Projections of the cost of cancer care in the United States: 2010-2020. *J Natl Cancer Inst,* 103**,** 117-28.

MARKOWITZ, S. D. & BERTAGNOLLI, M. M. 2009. Molecular origins of cancer: Molecular basis of colorectal cancer. *N Engl J Med,* 361**,** 2449-60.

MARTIN, G. S. 2004. The road to Src. *Oncogene,* 23**,** 7910-7.

MARTIN, M. A., BRYAN, T., RASHEED, S. & KHAN, A. S. 1981. Identification and cloning of endogenous retroviral sequences present in human DNA. *Proc Natl Acad Sci U S A,* 78**,** 4892-6.

MATSUDA, Y. & ICHIDA, T. 2009. Impact of hepatitis B virus X protein on the DNA damage response during hepatocarcinogenesis. *Med Mol Morphol,* 42**,** 138-42.

MATTOCK, H. 2012. *IARC Monographs on the evaluation of carcinogenic risks to humans* [Online]. Lyon, France: International Agency for Research on Cancer, World Health Organization. Available: http://monographs.iarc.fr/ENG/Classification/index.php [Accessed 26 January 2013 2013].

MCLAUGHLIN-DRUBIN, M. E. & MUNGER, K. 2008. Viruses associated with human cancer. *Biochim Biophys Acta,* 1782**,** 127-50.

MEDEMA, J. P. 2013. Cancer stem cells: the challenges ahead. *Nat Cell Biol,* 15**,** 338-44.

MEDZHITOV, R. 2008. Origin and physiological roles of inflammation. *Nature,* 454**,** 428-35.

MEEK, D. W. 2009. Tumour suppression by p53: a role for the DNA damage response? *Nat Rev Cancer,* 9**,** 714-23.

MEISLER, M. H. 2001. Evolutionarily conserved noncoding DNA in the human genome: how much and what for? *Genome Res,* 11**,** 1617-8.

MI, S., LEE, X., LI, X., VELDMAN, G. M., FINNERTY, H., RACIE, L., LAVALLIE, E., TANG, X. Y., EDOUARD, P., HOWES, S., KEITH, J. C., JR. & MCCOY, J. M. 2000. Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis. *Nature,* 403**,** 785-9.

MIKKERS, H. & BERNS, A. 2003. Retroviral insertional mutagenesis: tagging cancer pathways. *Adv Cancer Res,* 88**,** 53-99.

MILBRANDT, J. 1987. A nerve growth factor-induced gene encodes a possible transcriptional regulatory factor. *Science,* 238**,** 797-9.

MINISTRY OF HEALTH, S. 2010. *Cancer Screening: MOH Clinical Practice Guidelines 1/2010,* Singapore, Ministry of Health, Singapore.

MIRAGLIA, S., GODFREY, W., YIN, A. H., ATKINS, K., WARNKE, R., HOLDEN, J. T., BRAY, R. A., WALLER, E. K. & BUCK, D. W. 1997. A novel five-transmembrane hematopoietic stem cell antigen: isolation, characterization, and molecular cloning. *Blood,* 90**,** 5013-21.

MITSUI, K., TOKUZAWA, Y., ITOH, H., SEGAWA, K., MURAKAMI, M., TAKAHASHI, K., MARUYAMA, M., MAEDA, M. & YAMANAKA, S. 2003. The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell,* 113**,** 631-42.

MOIR, S., CHUN, T. W. & FAUCI, A. S. 2011. Pathogenic mechanisms of HIV disease. *Annu Rev Pathol,* 6**,** 223-48.

MOROZOV, V. A., DAO THI, V. L. & DENNER, J. 2013. The transmembrane protein of the human endogenous retrovirus--K (HERV-K) modulates cytokine release and gene expression. *PLoS One,* 8**,** e70399.

MORRISON, D. K. 2012. MAP kinase pathways. *Cold Spring Harb Perspect Biol,* 4.

MOTAVAF, M., SAFARI, S., SAFFARI JOURSHARI, M. & ALAVIAN, S. M. 2013. Hepatitis B virus-induced hepatocellular carcinoma: The role of the virus x protein. *Acta Virol,* 57**,** 389-96.

MOU, X., CHEN, L., LIU, F., LIN, J., DIAO, P., WANG, H., LI, Y., LIN, J., TENG, L. & XIANG, C. 2012. Prevalence of JC virus in Chinese patients with colorectal cancer. *PLoS One,* 7**,** e35900.

MOYES, D., GRIFFITHS, D. J. & VENABLES, P. J. 2007. Insertional polymorphisms: a new lease of life for endogenous retroviruses in human disease. *Trends Genet,* 23**,** 326-33.

MUIR, A., RUAN, Q. G., MARRON, M. P. & SHE, J. X. 1999. The IDDMK(1,2)22 retrovirus is not detectable in either mRNA or genomic DNA from patients with type 1 diabetes. *Diabetes,* 48**,** 219-22.

MULLINS, C. S. & LINNEBACHER, M. 2012. Human endogenous retroviruses and cancer: causality and therapeutic possibilities. *World J Gastroenterol,* 18**,** 6027-35.

MUNGER, K. & HOWLEY, P. M. 2002. Human papillomavirus immortalization and transformation functions. *Virus Res,* 89**,** 213-28.

NAGARAJ, S. H. & REVERTER, A. 2011. A Boolean-based systems biology approach to predict novel genes associated with cancer: Application to colorectal cancer. *BMC Syst Biol,* 5**,** 35.

NAGEL, S., MEYER, C., QUENTMEIER, H., KAUFMANN, M., DREXLER, H. G. & MACLEOD, R. A. 2008. MEF2C is activated by multiple mechanisms in a subset of T-acute lymphoblastic leukemia cell lines. *Leukemia,* 22**,** 600-7.

NAITO, T., OGASAWARA, H., KANEKO, H., HISHIKAWA, T., SEKIGAWA, I., HASHIMOTO, H. & MARUYAMA, N. 2003. Immune abnormalities induced by

human endogenous retroviral peptides: with reference to the pathogenesis of systemic lupus erythematosus. *J Clin Immunol,* 23**,** 371-6.

NATHANSON, N. 2008. The pathogenesis of poliomyelitis: what we don't know. *Adv Virus Res,* 71**,** 1-50.

NATIONAL REGISTRY OF DISEASES OFFICE 2010. Trends in cancer incidence in Singapore 1968 - 2007. *Singapore Cancer Registry Report No. 7.* Singapore: Singapore Cancer Registry.

NEIL, S. J., ZANG, T. & BIENIASZ, P. D. 2008. Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. *Nature,* 451**,** 425-30.

NELSON, A., DE SOYZA, A., PERRY, J. D., SUTCLIFFE, I. C. & CUMMINGS, S. P. 2012. Polymicrobial challenges to Koch's postulates: ecological lessons from the bacterial vaginosis and cystic fibrosis microbiomes. *Innate Immun,* 18**,** 774-83.

NEXO, B. A., CHRISTENSEN, T., FREDERIKSEN, J., MOLLER-LARSEN, A., OTURAI, A. B., VILLESEN, P., HANSEN, B., NISSEN, K. K., LASKA, M. J., PETERSEN, T. S., BONNESEN, S., HEDEMAND, A., WU, T., WANG, X., ZHANG, X., BRUDEK, T., MARIC, R., SONDERGAARD, H. B., SELLEBJERG, F., BRUSGAARD, K., KJELDBJERG, A. L., RASMUSSEN, H. B., NIELSEN, A. L., NYEGAARD, M., PETERSEN, T., BORGLUM, A. D. & PEDERSEN, F. S. 2011. The etiology of multiple sclerosis: genetic evidence for the involvement of the human endogenous retrovirus HERV-Fc1. *PLoS One,* 6**,** e16652.

NG, S. A. & LEE, C. 2011. Hepatitis B virus X gene and hepatocarcinogenesis. *J Gastroenterol,* 46**,** 974-90.

NGUYEN, L. V., VANNER, R., DIRKS, P. & EAVES, C. J. 2012. Cancer stem cells: an evolving concept. *Nat Rev Cancer,* 12**,** 133-43.

NIEHRS, C. 2012. The complex world of WNT receptor signalling. *Nat Rev Mol Cell Biol,* 13**,** 767-79.

NIIDA, A., HIROKO, T., KASAI, M., FURUKAWA, Y., NAKAMURA, Y., SUZUKI, Y., SUGANO, S. & AKIYAMA, T. 2004. DKK1, a negative regulator of Wnt signaling, is a target of the beta-catenin/TCF pathway. *Oncogene,* 23**,** 8520-6.

NIMROD, G., GLASER, F., STEINBERG, D., BEN-TAL, N. & PUPKO, T. 2005. In silico identification of functional regions in proteins. *Bioinformatics,* 21 Suppl 1**,** i328-37.

NING, B., DIAL, S., SUN, Y., WANG, J., YANG, J. & GUO, L. 2008. Systematic and simultaneous gene profiling of 84 drug-metabolizing genes in primary human hepatocytes. *J Biomol Screen,* 13**,** 194-201.

NISHIOKA, M., UENO, K., HAZAMA, S., OKADA, T., SAKAI, K., SUEHIRO, Y., OKAYAMA, N., HIRATA, H., OKA, M., IMAI, K., DAHIYA, R. & HINODA, Y. 2013. Possible involvement of Wnt11 in colorectal cancer progression. *Mol Carcinog,* 52**,** 207-17.

NISSEN, K. K., LASKA, M. J., HANSEN, B., TERKELSEN, T., VILLESEN, P., BAHRAMI, S., PETERSEN, T., PEDERSEN, F. S. & NEXO, B. A. 2013. Endogenous retroviruses and multiple sclerosis--new pieces to the puzzle. *BMC Neurol,* 13**,** 111.

NIV, Y., GOEL, A. & BOLAND, C. R. 2005. JC virus and colorectal cancer: a possible trigger in the chromosomal instability pathways. *Curr Opin Gastroenterol,* 21**,** 85-9.

NIV, Y., VILKIN, A., BRENNER, B., KENDEL, Y., MORGENSTERN, S. & LEVI, Z. 2010. hMLH1 promoter methylation and JC virus T antigen presence in the tumor tissue of colorectal cancer Israeli patients of different ethnic groups. *Eur J Gastroenterol Hepatol,* 22**,** 938-41.

NIWA, H., MIYAZAKI, J. & SMITH, A. G. 2000. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat Genet,* 24**,** 372-6.

NUSSE, R. & VARMUS, H. E. 1982. Many tumors induced by the mouse mammary tumor virus contain a provirus integrated in the same region of the host genome. *Cell,* 31**,** 99-109.

O'BRIEN, C. A., POLLETT, A., GALLINGER, S. & DICK, J. E. 2007. A human colon cancer cell capable of initiating tumour growth in immunodeficient mice. *Nature,* 445**,** 106-10.

O'CONNELL, C., O'BRIEN, S., NASH, W. G. & COHEN, M. 1984. ERV3, a full-length human endogenous provirus: chromosomal localization and evolutionary relationships. *Virology,* 138**,** 225-35.

O'DONOVAN, K. J., TOURTELLOTTE, W. G., MILLBRANDT, J. & BARABAN, J. M. 1999. The EGR family of transcription-regulatory factors: progress at the interface of molecular and systems neuroscience. *Trends Neurosci,* 22**,** 167-73.

ODEGAARD, A. O., KOH, W. P. & YUAN, J. M. 2013. Combined lifestyle factors and risk of incident colorectal cancer in a Chinese population. *Cancer Prev Res (Phila),* 6**,** 360-7.

OGINO, S. & GOEL, A. 2008. Molecular classification and correlates in colorectal cancer. *J Mol Diagn,* 10**,** 13-27.

OHTA, H., AOYAGI, K., FUKAYA, M., DANJOH, I., OHTA, A., ISOHATA, N., SAEKI, N., TANIGUCHI, H., SAKAMOTO, H., SHIMODA, T., TANI, T., YOSHIDA, T. & SASAKI, H. 2009. Cross talk between hedgehog and epithelial-mesenchymal transition pathways in gastric pit cells and in diffuse-type gastric cancers. *Br J Cancer,* 100**,** 389-98.

OKUMURA-NAKANISHI, S., SAITO, M., NIWA, H. & ISHIKAWA, F. 2005. Oct-3/4 and Sox2 regulate Oct-3/4 gene in embryonic stem cells. *J Biol Chem,* 280**,** 5307-17.

OLIVIER, M., HOLLSTEIN, M. & HAINAUT, P. 2010. TP53 mutations in human cancers: origins, consequences, and clinical use. *Cold Spring Harb Perspect Biol,* 2**,** a001008.

ONO, M., YASUNAGA, T., MIYATA, T. & USHIKUBO, H. 1986. Nucleotide sequence of human endogenous retrovirus genome related to the mouse mammary tumor virus genome. *J Virol,* 60**,** 589-98.

OUKO, L., ZIEGLER, T. R., GU, L. H., EISENBERG, L. M. & YANG, V. W. 2004. Wnt11 signaling promotes proliferation, transformation, and migration of IEC6 intestinal epithelial cells. *J Biol Chem,* 279**,** 26707-15.

OZEN, S. & BILGINER, Y. 2013. A clinical guide to autoinflammatory diseases: familial Mediterranean fever and next-of-kin. *Nat Rev Rheumatol*.

PAESSLER, S. & WALKER, D. H. 2012. Pathogenesis of the Viral Hemorrhagic Fevers. *Annu Rev Pathol*.

PALMER, T. D., ASHBY, W. J., LEWIS, J. D. & ZIJLSTRA, A. 2011. Targeting tumor cell motility to prevent metastasis. *Adv Drug Deliv Rev,* 63**,** 568-81.

PAN, G. & THOMSON, J. A. 2007. Nanog and transcriptional networks in embryonic stem cell pluripotency. *Cell Res,* 17**,** 42-9.

PANDUR, P., LASCHE, M., EISENBERG, L. M. & KUHL, M. 2002. Wnt-11 activation of a non-canonical Wnt signalling pathway is required for cardiogenesis. *Nature,* 418**,** 636-41.

PAPKOFF, J., BROWN, A. M. & VARMUS, H. E. 1987. The int-1 proto-oncogene products are glycoproteins that appear to enter the secretory pathway. *Mol Cell Biol,* 7**,** 3978-84.

PARADA, N. A., CENTER, D. M., KORNFELD, H., RODRIGUEZ, W. L., COOK, J., VALLEN, M. & CRUIKSHANK, W. W. 1998. Synergistic activation of CD4+ T cells by IL-16 and IL-2. *J Immunol,* 160**,** 2115-20.

PARKIN, D. M. 2006. The global health burden of infection-associated cancers in the year 2002. *Int J Cancer,* 118**,** 3030-44.

PARKIN, D. M., PISAM, P., MUNOZ, N. & FERLAY, J. 1999. The global health burden of infection associated cancers. *In:* NEWTON R, B. V., WEISS R (ed.) *Infections and Human Cancer.* New York: Cold Spring Harbor Laboratory Press.

PASSAM, F. H., SFIRIDAKI, A., PAPPA, C., KYRIAKOU, D., PETRELI, E., ROUSSOU, P. A. & ALEXANDRAKIS, M. G. 2008. Angiogenesis-related growth factors and cytokines in the serum of patients with B non-Hodgkin lymphoma; relation to clinical features and response to treatment. *Int J Lab Hematol,* 30**,** 17-25.

PASTRANA, E., SILVA-VARGAS, V. & DOETSCH, F. 2011. Eyes wide open: a critical review of sphere-formation as an assay for stem cells. *Cell Stem Cell,* 8**,** 486-98.

PERL, A., NAGY, G., KONCZ, A., GERGELY, P., FERNANDEZ, D., DOHERTY, E., TELARICO, T., BONILLA, E. & PHILLIPS, P. E. 2008. Molecular mimicry and immunomodulation by the HRES-1 endogenous retrovirus in SLE. *Autoimmunity,* 41**,** 287-97.

PERRON, H., HAMDANI, N., FAUCARD, R., LAJNEF, M., JAMAIN, S., DABAN-HUARD, C., SARRAZIN, S., LEGUEN, E., HOUENOU, J., DELAVEST, M., MOINS-TEISERENC, H., BENGOUFA, D., YOLKEN, R., MADEIRA, A., GARCIA-MONTOJO, M., GEHIN, N., BURGELIN, I., OLLAGNIER, G., BERNARD, C., DUMAINE, A., HENRION, A., GOMBERT, A., LE DUDAL, K., CHARRON, D., KRISHNAMOORTHY, R., TAMOUZA, R. & LEBOYER, M. 2012a. Molecular characteristics of Human Endogenous Retrovirus type-W in schizophrenia and bipolar disorder. *Transl Psychiatry,* 2**,** e201.

PERRON, H., HAMDANI, N., FAUCARD, R., LAJNEF, M., JAMAIN, S., DABAN-HUARD, C., SARRAZIN, S., LEGUEN, E., HOUENOU, J., DELAVEST, M., MOINS-TEISSERENC, H., BENGOUFA, D., YOLKEN, R., MADEIRA, A., GARCIA-MONTOJO, M., GEHIN, N., BURGELIN, I., OLLAGNIER, G., BERNARD, C., DUMAINE, A., HENRION, A., GOMBERT, A., LE DUDAL, K., CHARRON, D., KRISHNAMOORTHY, R., TAMOUZA, R. & LEBOYER, M. 2012b. Molecular characteristics of Human Endogenous Retrovirus type-W in schizophrenia and bipolar disorder. *Transl Psychiatry,* 2**,** e201.

PERRON, H., MEKAOUI, L., BERNARD, C., VEAS, F., STEFAS, I. & LEBOYER, M. 2008. Endogenous retrovirus type W GAG and envelope protein antigenemia in serum of schizophrenic patients. *Biol Psychiatry,* 64**,** 1019-23.

PETERSEN, T., MOLLER-LARSEN, A., ELLERMANN-ERIKSEN, S., THIEL, S. & CHRISTENSEN, T. 2012. Effects of interferon-beta therapy on elements in the antiviral immune response towards the human herpesviruses EBV, HSV, and VZV, and to the human endogenous retroviruses HERV-H and HERV-W in multiple sclerosis. *J Neuroimmunol,* 249**,** 105-8.

PHILIP, M., ROWLEY, D. A. & SCHREIBER, H. 2004. Inflammation as a tumor promoter in cancer induction. *Semin Cancer Biol,* 14**,** 433-9.

PIRES-DASILVA, A. & SOMMER, R. J. 2003. The evolution of signalling pathways in animal development. *Nat Rev Genet,* 4**,** 39-49.

PRAK, E. T. & KAZAZIAN, H. H., JR. 2000. Mobile elements and the human genome. *Nat Rev Genet,* 1**,** 134-44.

PRICE, E. J. & VENABLES, P. J. 1995. The etiopathogenesis of Sjogren's syndrome. *Semin Arthritis Rheum,* 25**,** 117-33.

QUERIDO, E., BLANCHETTE, P., YAN, Q., KAMURA, T., MORRISON, M., BOIVIN, D., KAELIN, W. G., CONAWAY, R. C., CONAWAY, J. W. & BRANTON, P. E. 2001. Degradation of p53 by adenovirus E4orf6 and E1B55K proteins occurs via a novel mechanism involving a Cullin-containing complex. *Genes Dev,* 15**,** 3104-17.

RACANIELLO, V. R. 2006. One hundred years of poliovirus pathogenesis. *Virology,* 344**,** 9-16.

RAMPIAS, T., BOUTATI, E., PECTASIDES, E., SASAKI, C., KOUNTOURAKIS, P., WEINBERGER, P. & PSYRRI, A. 2010. Activation of Wnt signaling pathway by human papillomavirus E6 and E7 oncogenes in HPV16-positive oropharyngeal squamous carcinoma cells. *Mol Cancer Res,* 8**,** 433-43.

RANDI, G., EDEFONTI, V., FERRARONI, M., LA VECCHIA, C. & DECARLI, A. 2010. Dietary patterns and the risk of colorectal cancer and adenomas. *Nutr Rev,* 68**,** 389-408.

RAULIN-CERCEAU, F., MAUREL, M. C. & SCHNEIDER, J. 1998. From Panspermia to Bioastronomy, the evolution of the hypothesis of universal life. *Orig Life Evol Biosph,* 28**,** 597-612.

REIS, B. S., JUNGBLUTH, A. A., FROSINA, D., HOLZ, M., RITTER, E., NAKAYAMA, E., ISHIDA, T., OBATA, Y., CARVER, B., SCHER, H., SCARDINO, P. T., SLOVIN, S., SUBUDHI, S. K., REUTER, V. E., SAVAGE, C., ALLISON, J. P., MELAMED, J., JAGER, E., RITTER, G., OLD, L. J. & GNJATIC, S. 2013. Prostate Cancer Progression Correlates with Increased Humoral Immune Response to a Human Endogenous Retrovirus GAG Protein. *Clin Cancer Res,* 19**,** 6112-25.

REMO, A., PANCIONE, M., ZANELLA, C. & VENDRAMINELLI, R. 2012. Molecular pathology of colorectal carcinoma. A systematic review centred on the new role of the pathologist. *Pathologica,* 104**,** 432-41.

RESH, M. D. 2005. Intracellular trafficking of HIV-1 Gag: how Gag interacts with cell membranes and makes viral particles. *AIDS Rev,* 7**,** 84-91.

REYNIER, F., VERJAT, T., TURREL, F., IMBERT, P. E., MAROTTE, H., MOUGIN, B. & MIOSSEC, P. 2009. Increase in human endogenous retrovirus HERV-K (HML-2) viral load in active rheumatoid arthritis. *Scand J Immunol,* 70**,** 295-9.

RICCI-VITIANI, L., LOMBARDI, D. G., PILOZZI, E., BIFFONI, M., TODARO, M., PESCHLE, C. & DE MARIA, R. 2007. Identification and expansion of human colon-cancer-initiating cells. *Nature,* 445**,** 111-5.

RIDER, P., CARMI, Y., GUTTMAN, O., BRAIMAN, A., COHEN, I., VORONOV, E., WHITE, M. R., DINARELLO, C. A. & APTE, R. N. 2011. IL-1alpha and IL-1beta recruit different myeloid cells and promote different stages of sterile inflammation. *J Immunol,* 187**,** 4835-43.

RIGGS, P. K., RHO, O. & DIGIOVANNI, J. 2000. Alteration of Egr-1 mRNA during multistage carcinogenesis in mouse skin. *Mol Carcinog,* 27**,** 247-51.

RILEY, T., SONTAG, E., CHEN, P. & LEVINE, A. 2008. Transcriptional control of human p53-regulated genes. *Nat Rev Mol Cell Biol,* 9**,** 402-12.

ROBBINS, E. W., TRAVANTY, E. A., YANG, K. & ICZKOWSKI, K. A. 2008. MAP kinase pathways and calcitonin influence CD44 alternate isoform expression in prostate cancer cells. *BMC Cancer,* 8**,** 260.

ROBERTS, M. L. & COOPER, N. R. 1998. Activation of a ras-MAPK-dependent pathway by Epstein-Barr virus latent membrane protein 1 is essential for cellular transformation. *Virology,* 240**,** 93-9.

RODRIGUEZ, L. G., WU, X. & GUAN, J. L. 2005. Wound-healing assay. *Methods Mol Biol,* 294**,** 23-9.

ROLLI, M., KOTLYAROV, A., SAKAMOTO, K. M., GAESTEL, M. & NEININGER, A. 1999. Stress-induced stimulation of early growth response gene-1 by p38/stress-activated protein kinase 2 is mediated by a cAMP-responsive promoter element in a MAPKAP kinase 2-independent manner. *J Biol Chem,* 274**,** 19559-64.

ROOSSINCK, M. J. 2011. The good viruses: viral mutualistic symbioses. *Nat Rev Microbiol,* 9**,** 99-108.

ROSENBERG, N. A. J., P. 1997. Retroviral pathogenesis. *In:* COFFIN, J. M., HUGHES,S.H. AND VARMUS,H.E (ed.) *Retroviruses. .* Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

ROUS, P. 1911. A SARCOMA OF THE FOWL TRANSMISSIBLE BY AN AGENT SEPARABLE FROM THE TUMOR CELLS. *J Exp Med,* 13**,** 397-411.

ROUS, P. 1983. Landmark article (JAMA 1911;56:198). Transmission of a malignant new growth by means of a cell-free filtrate. By Peyton Rous. *JAMA,* 250**,** 1445-9.

ROWE, H. M. & TRONO, D. 2011. Dynamic control of endogenous retroviruses during development. *Virology,* 411**,** 273-87.

RYAN, F. P. 2004. Human endogenous retroviruses in health and disease: a symbiotic perspective. *J R Soc Med,* 97**,** 560-5.

RYAN, G. B. & MAJNO, G. 1977. Acute inflammation. A review. *Am J Pathol,* 86**,** 183-276.

SAEGUSA, M., HASHIMURA, M., KUWATA, T., HAMANO, M., WATANABE, J., KAWAGUCHI, M. & OKAYASU, I. 2008. Transcription factor Egr1 acts as an upstream regulator of beta-catenin signalling through up-regulation of TCF4 and p300 expression during trans-differentiation of endometrial carcinoma cells. *J Pathol,* 216**,** 521-32.

SAHA, A., KAUL, R., MURAKAMI, M. & ROBERTSON, E. S. 2010. Tumor viruses and cancer biology: Modulating signaling pathways for therapeutic intervention. *Cancer Biol Ther,* 10**,** 961-78.

SAKAMOTO, T., OZAKI, K., FUJIO, K., KAJIKAWA, S. H., UESATO, S., WATANABE, K., TANIMURA, S., KOJI, T. & KOHNO, M. 2013. Blockade of the ERK pathway enhances the therapeutic efficacy of the histone deacetylase inhibitor MS-275 in human tumor xenograft models. *Biochem Biophys Res Commun,* 433**,** 456-62.

SANTONI, F. A., GUERRA, J. & LUBAN, J. 2012. HERV-H RNA is abundant in human embryonic stem cells and a precise marker for pluripotency. *Retrovirology,* 9**,** 111.

SARID, R. & GAO, S. J. 2011. Viruses and human cancer: from detection to causality. *Cancer Lett,* 305**,** 218-27.

SARNOW, P., HO, Y. S., WILLIAMS, J. & LEVINE, A. J. 1982. Adenovirus E1b-58kd tumor antigen and SV40 large tumor antigen are physically associated with the same 54 kd cellular protein in transformed cells. *Cell,* 28**,** 387-94.

SATO, Y., KAMURA, T., SHIRATA, N., MURATA, T., KUDOH, A., IWAHORI, S., NAKAYAMA, S., ISOMURA, H., NISHIYAMA, Y. & TSURUMI, T. 2009. Degradation of phosphorylated p53 by viral protein-ECS E3 ligase complex. *PLoS Pathog,* 5**,** e1000530.

SCHAEFFER, H. J. & WEBER, M. J. 1999. Mitogen-activated protein kinases: specific messages from ubiquitous messengers. *Mol Cell Biol,* 19**,** 2435-44.

SCHARNHORST, V., MENKE, A. L., ATTEMA, J., HANEVELD, J. K., RITECO, N., VAN STEENBRUGGE, G. J., VAN DER EB, A. J. & JOCHEMSEN, A. G. 2000. EGR-1

enhances tumor growth and modulates the effect of the Wilms' tumor 1 gene products on tumorigenicity. *Oncogene,* 19**,** 791-800.

SCHEFFNER, M., HUIBREGTSE, J. M., VIERSTRA, R. D. & HOWLEY, P. M. 1993. The HPV-16 E6 and E6-AP complex functions as a ubiquitin-protein ligase in the ubiquitination of p53. *Cell,* 75**,** 495-505.

SCHEFFNER, M., WERNESS, B. A., HUIBREGTSE, J. M., LEVINE, A. J. & HOWLEY, P. M. 1990. The E6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53. *Cell,* 63**,** 1129-36.

SCHEPERS, A. & CLEVERS, H. 2012. Wnt signaling, stem cells, and cancer of the gastrointestinal tract. *Cold Spring Harb Perspect Biol,* 4**,** a007989.

SCHUERWEGH, A. J., DOMBRECHT, E. J., STEVENS, W. J., VAN OFFEL, J. F., BRIDTS, C. H. & DE CLERCK, L. S. 2003. Influence of pro-inflammatory (IL-1 alpha, IL-6, TNF-alpha, IFN-gamma) and anti-inflammatory (IL-4) cytokines on chondrocyte function. *Osteoarthritis Cartilage,* 11**,** 681-7.

SCHULZ, W. A., STEINHOFF, C. & FLORL, A. R. 2006. Methylation of endogenous human retroelements in health and disease. *Curr Top Microbiol Immunol,* 310**,** 211-50.

SCHWIEGER, M., SCHULER, A., FORSTER, M., ENGELMANN, A., ARNOLD, M. A., DELWEL, R., VALK, P. J., LOHLER, J., SLANY, R. K., OLSON, E. N. & STOCKING, C. 2009. Homing and invasiveness of MLL/ENL leukemic cells is regulated by MEF2C. *Blood,* 114**,** 2476-88.

SEBOLT-LEOPOLD, J. S., DUDLEY, D. T., HERRERA, R., VAN BECELAERE, K., WILAND, A., GOWAN, R. C., TECLE, H., BARRETT, S. D., BRIDGES, A., PRZYBRANOWSKI, S., LEOPOLD, W. R. & SALTIEL, A. R. 1999. Blockade of the MAP kinase pathway suppresses growth of colon tumors in vivo. *Nat Med,* 5**,** 810-6.

SEGRE, J. A. 2013. What does it take to satisfy Koch's postulates two centuries later? Microbial genomics and Propionibacteria acnes. *J Invest Dermatol,* 133**,** 2141-2.

SEIFARTH, W., BAUST, C., SCHON, U., REICHERT, A., HEHLMANN, R. & LEIB-MOSCH, C. 2000. HERV-IP-T47D, a novel type C-related human endogenous retroviral sequence derived from T47D particles. *AIDS Res Hum Retroviruses,* 16**,** 471-80.

SEIFARTH, W., SKLADNY, H., KRIEG-SCHNEIDER, F., REICHERT, A., HEHLMANN, R. & LEIB-MOSCH, C. 1995. Retrovirus-like particles released from the human breast cancer cell line T47-D display type B- and C-related endogenous retroviral sequences. *J Virol,* 69**,** 6408-16.

SEMENOV, M. V., TAMAI, K., BROTT, B. K., KUHL, M., SOKOL, S. & HE, X. 2001. Head inducer Dickkopf-1 is a ligand for Wnt coreceptor LRP6. *Curr Biol,* 11**,** 951-61.

SENGUPTA, D., TANDON, R., VIEIRA, R. G., NDHLOVU, L. C., LOWN-HECHT, R., ORMSBY, C. E., LOH, L., JONES, R. B., GARRISON, K. E., MARTIN, J. N., YORK, V. A., SPOTTS, G., REYES-TERAN, G., OSTROWSKI, M. A., HECHT, F. M., DEEKS, S. G. & NIXON, D. F. 2011. Strong human endogenous retrovirus-specific T cell responses are associated with control of HIV-1 in chronic infection. *J Virol,* 85**,** 6977-85.

SERAFINO, A., BALESTRIERI, E., PIERIMARCHI, P., MATTEUCCI, C., MORONI, G., ORICCHIO, E., RASI, G., MASTINO, A., SPADAFORA, C., GARACI, E. & VALLEBONA, P. S. 2009. The activation of human endogenous retrovirus K (HERV-K) is implicated in melanoma cell malignant transformation. *Exp Cell Res,* 315**,** 849-62.

SGAMBATO, A. & CITTADINI, A. 2010. Inflammation and cancer: a multifaceted link. *Eur Rev Med Pharmacol Sci,* 14**,** 263-8.

SHARKEY, F. E. & FOGH, J. 1984. Considerations in the use of nude mice for cancer research. *Cancer Metastasis Rev,* 3**,** 341-60.

SHIH, A., COUTAVAS, E. E. & RUSH, M. G. 1991. Evolutionary implications of primate endogenous retroviruses. *Virology,* 182**,** 495-502.

SHIH, T. Y. & WEEKS, M. O. 1984. Oncogenes and cancer: the p21 ras genes. *Cancer Invest,* 2**,** 109-23.

SHIH, T. Y., WEEKS, M. O., YOUNG, H. A. & SCHOLNICK, E. M. 1979. Identification of a sarcoma virus-coded phosphoprotein in nonproducer cells transformed by Kirsten or Harvey murine sarcoma virus. *Virology,* 96**,** 64-79.

SHMELKOV, S. V., BUTLER, J. M., HOOPER, A. T., HORMIGO, A., KUSHNER, J., MILDE, T., ST CLAIR, R., BALJEVIC, M., WHITE, I., JIN, D. K., CHADBURN, A., MURPHY, A. J., VALENZUELA, D. M., GALE, N. W., THURSTON, G., YANCOPOULOS, G. D., D'ANGELICA, M., KEMENY, N., LYDEN, D. & RAFII, S. 2008. CD133 expression is not restricted to stem cells, and both CD133+ and CD133- metastatic colon cancer cells initiate tumors. *J Clin Invest,* 118**,** 2111-20.

SIDEMAN, S. 2005. Preface: the cellular communications maze. *Ann N Y Acad Sci,* 1047**,** xi-xxiv.

SINGH, S., KAYE, S., GORE, M. E., MCCLURE, M. O. & BUNKER, C. B. 2009. The role of human endogenous retroviruses in melanoma. *Br J Dermatol,* 161**,** 1225-31.

SINGH, S. K., HAWKINS, C., CLARKE, I. D., SQUIRE, J. A., BAYANI, J., HIDE, T., HENKELMAN, R. M., CUSIMANO, M. D. & DIRKS, P. B. 2004. Identification of human brain tumour initiating cells. *Nature,* 432**,** 396-401.

SLATTERY, M. L., LUNDGREEN, A., BONDURANT, K. L. & WOLFF, R. K. 2011. Tumor necrosis factor-related genes and colon and rectal cancer. *Int J Mol Epidemiol Genet,* 2**,** 328-38.

SLATTERY, M. L., LUNDGREEN, A. & WOLFF, R. K. 2012. MAP kinase genes and colon and rectal cancer. *Carcinogenesis,* 33**,** 2398-408.

SMITH, D. J. 2013. Microbes in the upper atmosphere and unique opportunities for astrobiology research. *Astrobiology,* 13**,** 981-90.

SOHN, K. C., SHI, G., JANG, S., CHOI, D. K., LEE, Y., YOON, T. J., PARK, H., HWANG, C., KIM, H. J., SEO, Y. J., LEE, J. H., PARK, J. K. & KIM, C. D. 2009. Pitx2, a beta-catenin-regulated transcription factor, regulates the differentiation of outer root sheath cells cultured in vitro. *J Dermatol Sci,* 54**,** 6-11.

SOLYOM, S. & KAZAZIAN, H. H., JR. 2012. Mobile elements in the human genome: implications for disease. *Genome Med,* 4**,** 12.

SONGOK, E. M., LUO, M., LIANG, B., MCLAREN, P., KAEFER, N., APIDI, W., BOUCHER, G., KIMANI, J., WACHIHI, C., SEKALY, R., FOWKE, K., BALL, B. T. & PLUMMER, F. A. 2012. Microarray analysis of HIV resistant female sex workers reveal a gene expression signature pattern reminiscent of a lowered immune activation state. *PLoS One,* 7**,** e30048.

STAUFFER, Y., MARGUERAT, S., MEYLAN, F., UCLA, C., SUTKOWSKI, N., HUBER, B., PELET, T. & CONRAD, B. 2001. Interferon-alpha-induced endogenous superantigen. a model linking environment and autoimmunity. *Immunity,* 15**,** 591-601.

STEELE, P. E., RABSON, A. B., BRYAN, T. & MARTIN, M. A. 1984. Distinctive termini characterize two families of human endogenous retroviral sequences. *Science,* 225**,** 943-7.

STOYE, J. P. & COFFIN, J. M. 2000. A provirus put to work. *Nature,* 403**,** 715, 717.

SU, Y. J., LAI, H. M., CHANG, Y. W., CHEN, G. Y. & LEE, J. L. 2011. Direct reprogramming of stem cell properties in colon cancer cells by CD44. *EMBO J,* 30**,** 3186-99.

SUETSUGU, A., NAGAKI, M., AOKI, H., MOTOHASHI, T., KUNISADA, T. & MORIWAKI, H. 2006. Characterization of CD133+ hepatocellular carcinoma cells as cancer stem/progenitor cells. *Biochem Biophys Res Commun,* 351**,** 820-4.

SUGIMURA, R., HE, X. C., VENKATRAMAN, A., ARAI, F., BOX, A., SEMERAD, C., HAUG, J. S., PENG, L., ZHONG, X. B., SUDA, T. & LI, L. 2012. Noncanonical Wnt signaling maintains hematopoietic stem cells in the niche. *Cell,* 150**,** 351-65.

SUGIMURA, R. & LI, L. 2010. Noncanonical Wnt signaling in vertebrate development, stem cells, and diseases. *Birth Defects Res C Embryo Today,* 90**,** 243-56.

SVAREN, J., EHRIG, T., ABDULKADIR, S. A., EHRENGRUBER, M. U., WATSON, M. A. & MILBRANDT, J. 2000. EGR1 target genes in prostate carcinoma cells identified by microarray analysis. *J Biol Chem,* 275**,** 38524-31.

SVERDLOV, E. 2005. Retrovirus, their domesticated relatives and other retroinvaders: Potential genetic and epigenetic mediators of phenotypic variation. *Retroviruses and primate genome evolution.* 1 ed. USA: Landes Bioscience.

TAI, A. K., O'REILLY, E. J., ALROY, K. A., SIMON, K. C., MUNGER, K. L., HUBER, B. T. & ASCHERIO, A. 2008. Human endogenous retrovirus-K18 Env as a risk factor in multiple sclerosis. *Mult Scler,* 14**,** 1175-80.

TAKAYAMA, T., MIYANISHI, K., HAYASHI, T., SATO, Y. & NIITSU, Y. 2006. Colorectal cancer: genetics of development and metastasis. *J Gastroenterol,* 41**,** 185-92.

TAKEUCHI, H. & MATANO, T. 2008. Host factors involved in resistance to retroviral infection. *Microbiol Immunol,* 52**,** 318-25.

TANDON, R., SENGUPTA, D., NDHLOVU, L. C., VIEIRA, R. G., JONES, R. B., YORK, V. A., VIEIRA, V. A., SHARP, E. R., WIZNIA, A. A., OSTROWSKI, M. A., ROSENBERG, M. G. & NIXON, D. F. 2011. Identification of human endogenous retrovirus-specific T cell responses in vertically HIV-1-infected subjects. *J Virol,* 85**,** 11526-31.

TANIMURA, S., ASATO, K., FUJISHIRO, S. H. & KOHNO, M. 2003. Specific blockade of the ERK pathway inhibits the invasiveness of tumor cells: down-regulation of matrix metalloproteinase-3/-9/-14 and CD44. *Biochem Biophys Res Commun,* 304**,** 801-6.

TARCIC, G., AVRAHAM, R., PINES, G., AMIT, I., SHAY, T., LU, Y., ZWANG, Y., KATZ, M., BEN-CHETRIT, N., JACOB-HIRSCH, J., VIRGILIO, L., RECHAVI, G., MAVROTHALASSITIS, G., MILLS, G. B., DOMANY, E. & YARDEN, Y. 2012. EGR1 and the ERK-ERF axis drive mammary cell migration in response to EGF. *FASEB J,* 26**,** 1582-92.

TEO, M. C. & SOO, K. C. 2013. Cancer trends and incidences in Singapore. *Jpn J Clin Oncol,* 43**,** 219-24.

TERZIC, J., GRIVENNIKOV, S., KARIN, E. & KARIN, M. 2010. Inflammation and colon cancer. *Gastroenterology,* 138**,** 2101-2114 e5.

THEOBALD, D. L. 2010. A formal test of the theory of universal common ancestry. *Nature,* 465**,** 219-22.

THEODOROPOULOS, G., PANOUSSOPOULOS, D., PAPACONSTANTINOU, I., GAZOULI, M., PERDIKI, M., BRAMIS, J. & LAZARIS, A. 2005. Assessment of JC polyoma virus in colon neoplasms. *Dis Colon Rectum,* 48**,** 86-91.

THIEL, G. & CIBELLI, G. 2002. Regulation of life and death by the zinc finger transcription factor Egr-1. *J Cell Physiol,* 193**,** 287-92.

TODARO, M., ALEA, M. P., DI STEFANO, A. B., CAMMARERI, P., VERMEULEN, L., IOVINO, F., TRIPODO, C., RUSSO, A., GULOTTA, G., MEDEMA, J. P. & STASSI, G. 2007. Colon cancer stem cells dictate tumor growth and resist cell death by production of interleukin-4. *Cell Stem Cell,* 1**,** 389-402.

TOM, B. H., RUTZKY, L. P., JAKSTYS, M. M., OYASU, R., KAYE, C. I. & KAHAN, B. D. 1976. Human colonic adenocarcinoma cells. I. Establishment and description of a new line. *In Vitro,* 12**,** 180-91.

TONARY, A. M., MACDONALD, E. A., FAUGHT, W., SENTERMAN, M. K. & VANDERHYDEN, B. C. 2000. Lack of expression of c-KIT in ovarian cancers is associated with poor prognosis. *Int J Cancer,* 89**,** 242-50.

TOWERS, G. J. 2005. Control of viral infectivity by tripartite motif proteins. *Hum Gene Ther,* 16**,** 1125-32.

TRISTEM, M. 2000. Identification and characterization of novel human endogenous retrovirus families by phylogenetic screening of the human genome mapping project database. *J Virol,* 74**,** 3715-30.

TSUTSUI, S., YASUDA, K., SUZUKI, K., TAKEUCHI, H., NISHIZAKI, T., HIGASHI, H. & ERA, S. 2006. A loss of c-kit expression is associated with an advanced stage and poor prognosis in breast cancer. *Br J Cancer,* 94**,** 1874-8.

TURNER, G., BARBULESCU, M., SU, M., JENSEN-SEAMAN, M. I., KIDD, K. K. & LENZ, J. 2001. Insertional polymorphisms of full-length endogenous retroviruses in humans. *Curr Biol,* 11**,** 1531-5.

UREN, A. G., KOOL, J., BERNS, A. & VAN LOHUIZEN, M. 2005. Retroviral insertional mutagenesis: past, present and future. *Oncogene,* 24**,** 7656-72.

URNOVITZ, H. B. & MURPHY, W. H. 1996. Human endogenous retroviruses: nature, occurrence, and clinical implications in human disease. *Clin Microbiol Rev,* 9**,** 72-99.

VADLAMUDI, U., ESPINOZA, H. M., GANGA, M., MARTIN, D. M., LIU, X., ENGELHARDT, J. F. & AMENDT, B. A. 2005. PITX2, beta-catenin and LEF-1 interact to synergistically regulate the LEF-1 promoter. *J Cell Sci,* 118**,** 1129-37.

VAN GRIEKEN, N. C., AOYMA, T., CHAMBERS, P. A., BOTTOMLEY, D., WARD, L. C., INAM, I., BUFFART, T. E., DAS, K., LIM, T., PANG, B., ZHANG, S. L., TAN, I. B., CARVALHO, B., HEIDEMAN, D. A., MIYAGI, Y., KAMEDA, Y., ARAI, T., MEIJER, G. A., TSUBURAYA, A., TAN, P., YOSHIKAWA, T. & GRABSCH, H. I. 2013. KRAS and BRAF mutations are rare and related to DNA mismatch repair deficiency in gastric cancer from the East and the West: results from a large international multicentre study. *Br J Cancer,* 108**,** 1495-501.

VANDER HEIDEN, M. G., CANTLEY, L. C. & THOMPSON, C. B. 2009. Understanding the Warburg effect: the metabolic requirements of cell proliferation. *Science,* 324**,** 1029-33.

VARELA, M., SPENCER, T. E., PALMARINI, M. & ARNAUD, F. 2009. Friendly viruses: the special relationship between endogenous retroviruses and their host. *Ann N Y Acad Sci,* 1178**,** 157-72.

VARMUS, H. 1988. Retroviruses. *Science,* 240**,** 1427-35.

VENDRAMINI-COSTA, D. B. & CARVALHO, J. E. 2012. Molecular link mechanisms between inflammation and cancer. *Curr Pharm Des,* 18**,** 3831-52.

VENKATESH, B., KIRKNESS, E. F., LOH, Y. H., HALPERN, A. L., LEE, A. P., JOHNSON, J., DANDONA, N., VISWANATHAN, L. D., TAY, A., VENTER, J. C., STRAUSBERG, R. L. & BRENNER, S. 2006. Ancient noncoding elements conserved in the human genome. *Science,* 314**,** 1892.

VERMEULEN, L., DE SOUSA, E. M. F., VAN DER HEIJDEN, M., CAMERON, K., DE JONG, J. H., BOROVSKI, T., TUYNMAN, J. B., TODARO, M., MERZ, C., RODERMOND, H., SPRICK, M. R., KEMPER, K., RICHEL, D. J., STASSI, G. & MEDEMA, J. P. 2010. Wnt activity defines colon cancer stem cells and is regulated by the microenvironment. *Nat Cell Biol,* 12**,** 468-76.

VIJAYARAGAVAN, K., SZABO, E., BOSSE, M., RAMOS-MEJIA, V., MOON, R. T. & BHATIA, M. 2009. Noncanonical Wnt signaling orchestrates early developmental events toward hematopoietic cell fate from human embryonic stem cells. *Cell Stem Cell,* 4**,** 248-62.

VILKIN, A. & NIV, Y. 2011. Association between hMLH1 hypermethylation and JC virus (JCV) infection in human colorectal cancer (CRC). *Clin Epigenetics,* 2**,** 1-5.

VILLARREAL, L. P. 2009. The source of self: genetic parasites and the origin of adaptive immunity. *Ann N Y Acad Sci,* 1178**,** 194-232.

VOGT, P. K. 2012. Retroviral oncogenes: a historical primer. *Nat Rev Cancer,* 12**,** 639-48.

VOISSET, C., WEISS, R. A. & GRIFFITHS, D. J. 2008. Human RNA "rumor" viruses: the search for novel human retroviruses in chronic disease. *Microbiol Mol Biol Rev,* 72**,** 157-96, table of contents.

VORONOV, E., SHOUVAL, D. S., KRELIN, Y., CAGNANO, E., BENHARROCH, D., IWAKURA, Y., DINARELLO, C. A. & APTE, R. N. 2003. IL-1 is required for tumor invasiveness and angiogenesis. *Proc Natl Acad Sci U S A,* 100**,** 2645-50.

VOUSDEN, K. H. & LANE, D. P. 2007. p53 in health and disease. *Nat Rev Mol Cell Biol,* 8**,** 275-83.

WALKER, L., LEVINE, H. & JUCKER, M. 2006. Koch's postulates and infectious proteins. *Acta Neuropathol,* 112**,** 1-4.

WANG-JOHANNING, F., FROST, A. R., JIAN, B., AZEROU, R., LU, D. W., CHEN, D. T. & JOHANNING, G. L. 2003. Detecting the expression of human endogenous retrovirus E envelope transcripts in human prostate adenocarcinoma. *Cancer,* 98**,** 187-97.

WANG-JOHANNING, F., RADVANYI, L., RYCAJ, K., PLUMMER, J. B., YAN, P., SASTRY, K. J., PIYATHILAKE, C. J., HUNT, K. K. & JOHANNING, G. L. 2008. Human endogenous retrovirus K triggers an antigen-specific immune response in breast cancer patients. *Cancer Res,* 68**,** 5869-77.

WANG, A. L. & WANG, C. C. 1991. Viruses of the protozoa. *Annu Rev Microbiol,* 45**,** 251-63.

WANG, H. Y., LIU, T. & MALBON, C. C. 2006. Structure-function analysis of Frizzleds. *Cell Signal,* 18**,** 934-41.

WANG, Q., ZHOU, Y., RYCHAHOU, P., LIU, C., WEISS, H. L. & EVERS, B. M. 2013. NFAT5 represses canonical Wnt signaling via inhibition of beta-catenin acetylation and participates in regulating intestinal cell differentiation. *Cell Death Dis,* 4**,** e671.

WANG, Y. K., ZHU, Y. L., QIU, F. M., ZHANG, T., CHEN, Z. G., ZHENG, S. & HUANG, J. 2010. Activation of Akt and MAPK pathways enhances the tumorigenicity of CD133+ primary colon cancer cells. *Carcinogenesis,* 31**,** 1376-80.

WANG, Z., LI, Y., LIU, E. T. & YU, Q. 2004. Susceptibility to cell death induced by blockade of MAPK pathway in human colorectal cancer cells carrying Ras mutations is dependent on p53 status. *Biochem Biophys Res Commun,* 322**,** 609-13.

WARBURG, O. 1956. On the origin of cancer cells. *Science,* 123**,** 309-14.

WARREN, J. L., YABROFF, K. R., MEEKINS, A., TOPOR, M., LAMONT, E. B. & BROWN, M. L. 2008. Evaluation of trends in the cost of initial cancer treatment. *J Natl Cancer Inst,* 100**,** 888-97.

WEG-REMERS, S., ANDERS, M., VON LAMPE, B., RIECKEN, E. O., SCHUDER, G., FEIFEL, G., ZEITZ, M. & STALLMACH, A. 1998. Decreased expression of CD44 splicing variants in advanced colorectal carcinomas. *Eur J Cancer,* 34**,** 1607-11.

WEG-REMERS, S., PONTA, H., HERRLICH, P. & KONIG, H. 2001. Regulation of alternative pre-mRNA splicing by the ERK MAP-kinase pathway. *EMBO J,* 20**,** 4194-203.

WEISS, R. A. 1996. Retrovirus classification and cell interactions. *J Antimicrob Chemother,* 37 Suppl B**,** 1-11.

WENTZENSEN, N., COY, J. F., KNAEBEL, H. P., LINNEBACHER, M., WILZ, B., GEBERT, J. & VON KNEBEL DOEBERITZ, M. 2007. Expression of an endogenous retroviral sequence from the HERV-H group in gastrointestinal cancers. *Int J Cancer,* 121**,** 1417-23.

WENTZENSEN, N., WILZ, B., FINDEISEN, P., WAGNER, R., DIPPOLD, W., VON KNEBEL DOEBERITZ, M. & GEBERT, J. 2004. Identification of differentially expressed genes in colorectal adenoma compared to normal tissue by suppression subtractive hybridization. *Int J Oncol,* 24**,** 987-94.

WERNECK, M. B., HOTTZ, E., BOZZA, P. T. & VIOLA, J. P. 2012. Cyclosporin A inhibits colon cancer cell growth independently of the calcineurin pathway. *Cell Cycle,* 11**,** 3997-4008.

WERNER, T., BRACK-WERNER, R., LEIB-MOSCH, C., BACKHAUS, H., ERFLE, V. & HEHLMANN, R. 1990. S71 is a phylogenetically distinct human endogenous retroviral element with structural and sequence homology to simian sarcoma virus (SSV). *Virology,* 174**,** 225-38.

WICKRAMASINGHE, N. C. & TREVORS, J. T. 2013. Non-terrestrial origin of life: a transformative research paradigm shift. *Theory Biosci,* 132**,** 133-7.

WILKINSON, D. A., MAGER, D. L. & LEONG, J. A. C. 1994. Endogenous human retroviruses. *In:* LEVY, J. A. (ed.) *The Retroviridae* New York: Plenum.

WILSON, J. M., COLETTA, P. L., CUTHBERT, R. J., SCOTT, N., MACLENNAN, K., HAWCROFT, G., LENG, L., LUBETSKY, J. B., JIN, K. K., LOLIS, E., MEDINA, F., BRIEVA, J. A., POULSOM, R., MARKHAM, A. F., BUCALA, R. & HULL, M. A. 2005. Macrophage migration inhibitory factor promotes intestinal tumorigenesis. *Gastroenterology,* 129**,** 1485-503.

WOESE, C. 1998. The universal ancestor. *Proc Natl Acad Sci U S A,* 95**,** 6854-9.

WOMMACK, K. E. & COLWELL, R. R. 2000. Virioplankton: viruses in aquatic ecosystems. *Microbiol Mol Biol Rev,* 64**,** 69-114.

WONG, M. T. & EU, K. W. 2007. Rise of colorectal cancer in Singapore: an epidemiological review. *ANZ J Surg,* 77**,** 446-9.

WOO, W. H., LEONG, S. M. & KOAY, E. S. The distribution of the endogenous retrovirus HERV-K113 among adolescents. Joint Conference of HGM 2013 and 21st International Congress of Genetics, 2013 Singapore.

WOOLFE, A., GOODSON, M., GOODE, D. K., SNELL, P., MCEWEN, G. K., VAVOURI, T., SMITH, S. F., NORTH, P., CALLAWAY, H., KELLY, K., WALTER, K., ABNIZOVA, I., GILKS, W., EDWARDS, Y. J., COOKE, J. E. & ELGAR, G. 2005. Highly conserved non-coding sequences are associated with vertebrate development. *PLoS Biol,* 3**,** e7.

WORTHLEY, D. L., WHITEHALL, V. L., SPRING, K. J. & LEGGETT, B. A. 2007. Colorectal carcinogenesis: road maps to cancer. *World J Gastroenterol,* 13**,** 3784-91.

WU, J., KUBOTA, J., HIRAYAMA, J., NAGAI, Y., NISHINA, S., YOKOI, T., ASAOKA, Y., SEO, J., SHIMIZU, N., KAJIHO, H., WATANABE, T., AZUMA, N., KATADA, T. & NISHINA, H. 2010. p38 Mitogen-activated protein kinase controls a switch between cardiomyocyte and neuronal commitment of murine embryonic stem cells by activating myocyte enhancer factor 2C-dependent bone morphogenetic protein 2 transcription. *Stem Cells Dev,* 19**,** 1723-34.

XU, W., CHOU, C. L., SUN, H., FUJINO, H., CHEN, Q. M. & REGAN, J. W. 2008. FP prostanoid receptor-mediated induction of the expression of early growth response factor-1 by activation of a Ras/Raf/mitogen-activated protein kinase signaling cascade. *Mol Pharmacol,* 73**,** 111-8.

YANG, H. S., MATTHEWS, C. P., CLAIR, T., WANG, Q., BAKER, A. R., LI, C. C., TAN, T. H. & COLBURN, N. H. 2006. Tumorigenesis suppressor Pdcd4 down-regulates mitogen-activated protein kinase kinase kinase kinase 1 expression to suppress colon carcinoma cell invasion. *Mol Cell Biol,* 26**,** 1297-306.

YANG, K., TANG, Y., HABERMEHL, G. K. & ICZKOWSKI, K. A. 2010. Stable alterations of CD44 isoform expression in prostate cancer cells decrease invasion and growth and alter ligand binding and chemosensitivity. *BMC Cancer,* 10**,** 16.

YEE, K. S. & VOUSDEN, K. H. 2005. Complicating the complexity of p53. *Carcinogenesis,* 26**,** 1317-22.

YEH, T. C., MARSH, V., BERNAT, B. A., BALLARD, J., COLWELL, H., EVANS, R. J., PARRY, J., SMITH, D., BRANDHUBER, B. J., GROSS, S., MARLOW, A., HURLEY, B., LYSSIKATOS, J., LEE, P. A., WINKLER, J. D., KOCH, K. & WALLACE, E. 2007. Biological characterization of ARRY-142886 (AZD6244), a potent, highly selective mitogen-activated protein kinase kinase 1/2 inhibitor. *Clin Cancer Res,* 13**,** 1576-83.

YIN, A. H., MIRAGLIA, S., ZANJANI, E. D., ALMEIDA-PORADA, G., OGAWA, M., LEARY, A. G., OLWEUS, J., KEARNEY, J. & BUCK, D. W. 1997. AC133, a novel marker for human hematopoietic stem and progenitor cells. *Blood,* 90**,** 5002-12.

YOUNG, G. R., STOYE, J. P. & KASSIOTIS, G. 2013. Are human endogenous retroviruses pathogenic? An approach to testing the hypothesis. *Bioessays,* 35**,** 794-803.

YU, G., SHEN, F. S., STURCH, S., AQUINO, A., GLAZER, R. I. & FELSTED, R. L. 1995. Regulation of HIV-1 gag protein subcellular targeting by protein kinase C. *J Biol Chem,* 270**,** 4792-6.

YU, H. L., ZHAO, Z. K. & ZHU, F. 2013. The role of human endogenous retroviral long terminal repeat sequences in human cancer (Review). *Int J Mol Med,* 32**,** 755-62.

YU, Q., KONIG, R., PILLAI, S., CHILES, K., KEARNEY, M., PALMER, S., RICHMAN, D., COFFIN, J. M. & LANDAU, N. R. 2004. Single-strand specificity of APOBEC3G accounts for minus-strand deamination of the HIV genome. *Nat Struct Mol Biol,* 11**,** 435-42.

YUN, J., JOHNSON, J. L., HANIGAN, C. L. & LOCASALE, J. W. 2012. Interactions between epigenetics and metabolism in cancers. *Front Oncol,* 2**,** 163.

ZAKUT, R., PERLIS, R., ELIYAHU, S., YARDEN, Y., GIVOL, D., LYMAN, S. D. & HALABAN, R. 1993. KIT ligand (mast cell growth factor) inhibits the growth of KIT-expressing melanoma cells. *Oncogene,* 8**,** 2221-9.

ZHANG, M., LIU, Y., FENG, H., BIAN, X., ZHAO, W., YANG, Z., GU, B., LI, Z. & LIU, Y. 2013a. CD133 affects the invasive ability of HCT116 cells by regulating TIMP-2. *Am J Pathol,* 182**,** 565-76.

ZHANG, P., CAI, Y., SOOFI, A. & DRESSLER, G. R. 2012. Activation of Wnt11 by transforming growth factor-beta drives mesenchymal gene expression

through non-canonical Wnt protein signaling in renal epithelial cells. *J Biol Chem,* 287**,** 21290-302.

ZHANG, X., ZHANG, Z., ZHENG, B., HE, Z., WINBERG, G. & ERNBERG, I. 2013b. An update on viral association of human cancers. *Arch Virol,* 158**,** 1433-43.

ZHAO, J., RYCAJ, K., GENG, S., LI, M., PLUMMER, J. B., YIN, B., LIU, H., XU, X., ZHANG, Y., YAN, Y., GLYNN, S. A., DORSEY, T. H., AMBS, S., JOHANNING, G. L., GU, L. & WANG-JOHANNING, F. 2011. Expression of Human Endogenous Retrovirus Type K Envelope Protein is a Novel Candidate Prognostic Marker for Human Breast Cancer. *Genes Cancer,* 2**,** 914-22.

ZHENG, X., BAKER, H., HANCOCK, W. S., FAWAZ, F., MCCAMAN, M. & PUNGOR, E., JR. 2006. Proteomic analysis for the assessment of different lots of fetal bovine serum as a raw material for cell culture. Part IV. Application of proteomics to the manufacture of biological drugs. *Biotechnol Prog,* 22**,** 1294-300.

ZHOU, G., LEE, S. C., YAO, Z. & TAN, T. H. 1999. Hematopoietic progenitor kinase 1 is a component of transforming growth factor beta-induced c-Jun N-terminal kinase signaling cascade. *J Biol Chem,* 274**,** 13133-8.

ZOLLER, M. 2011. CD44: can a cancer-initiating cell profit from an abundantly expressed molecule? *Nat Rev Cancer,* 11**,** 254-67.

ZUR HAUSEN, H. 1991. Viruses in human cancers. *Science,* 254**,** 1167-73.

ZUR HAUSEN, H. 2009. The search for infectious causes of human cancers: where and why. *Virology,* 392**,** 1-10.

ZWANG, Y., OREN, M. & YARDEN, Y. 2012. Consistency test of the cell cycle: roles for p53 and EGR1. *Cancer Res,* 72**,** 1051-4.

ZWANG, Y., SAS-CHEN, A., DRIER, Y., SHAY, T., AVRAHAM, R., LAURIOLA, M., SHEMA, E., LIDOR-NILI, E., JACOB-HIRSCH, J., AMARIGLIO, N., LU, Y., MILLS, G. B., RECHAVI, G., OREN, M., DOMANY, E. & YARDEN, Y. 2011. Two phases of mitogenic signaling unveil roles for p53 and EGR1 in elimination of inconsistent growth signals. *Mol Cell,* 42**,** 524-35.